

C;Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 30-Jun-2001

C/Accession: J07619
R/Takahashi, T.; Sugishita, Y.; Nojiri, T.; Shinizu, T.; Yao, A.; Kingawa, K.; Harada, Biochem. Biophys. Res. Commun. 281, 1057-1062, 2001
A/Title: Cloning of hypoxia-inducible factor 1 alpha cDNA from chick embryonic ventricle
A/Reference number: J07619; MUID:21134360; PMID:11237772
A/Contents: Embryonic ventricular myocytes
A/Accession: J07619
A/Molecule type: mRNA
A/Residues: 1-811 <TRAK>
A/Cross-references: DDBJ:AB013746
A/Comment: This factor belongs to the basic helix-loop-helix-Per/ARNT/Sim (HLH-PAS) family of cells, and in redox stimuli.
C/Genetics: hif-1alpha
C/Keywords: embryo; transcription factor
F/106-156/Domain: Per/ARNT/Sim, ligand binding, dimerization #status predicted <PAS1>
F/249-299/Domain: Per/ARNT/Sim, ligand binding, dimerization #status predicted <PAS2>
F/762-811/Domain: conserved carboxy-terminal transactivation element #status predicted <F/767-768/Region: conserved dileucine repeat, important for oxygen-dependent degradation

Query Match 30.7%; Score 503; DB 2; Length 811;
Best Local Similarity 49.8%; Pred. No. 8e-36;
Matches 119; Conservative 20; Mismatches 72; Indels 28; Gaps 5;

QY 9 RSNTELRKERSDARSRQETEVLYQALATLPFARGVSAHLDKASIMRLTISTLRMR 68
DB 12 RISERREKSRDARSRKSESEVFYELAHQLPLPHVSSHLDDASVWRLTISTLRMR 71
QY 69 LCAAGCKRGAT-----GRLPEPGGFRHGTNRGRHGLPVGKCOQAPGQSVLDCSS 123
DB 72 LLDGGELETEANMEKELNCFYKALDGFVWVLTSDG-----DMVYMS 113
QY 124 LIHNPTPG-TNFSLELIGHSIFDFIHPDOELDALTPRPNLSSKYLEAFTERHFLRM 182
DB 114 ENVAKCKMGLTQF--DLTGHSVDFPTHPDHEHEMLTHRNGPYKCKEQNTERSFELRM 171
QY 183 KSTLTSRGRITNLKAAITKYLHCSGMRAKYKRPNOTSPAGSPRSEPPLOCLVICEAIP 241
DB 172 KCTLTSRGRITNLKAAITKYLHCSGMRAKYKRPNOTSPAGSPRSEPPLOCLVICEAIP 228

RESULT 3
138972
hypoxia-inducible factor 1 alpha - human
N/Alternate names: ARNT interacting protein
C/Species: Homo sapiens (man)
C/Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 09-Jul-2004
C/Accession: I38972; G01875
R/Mang, G.L.; Jiang, B.H.; Rue, E.A.; Semenza, G.L. Proc. Natl. Acad. Sci. U.S.A. 92, 5510-5514, 1995
A/Title: Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulate
A/Reference number: I38972; MUID:95296340; PMID:7539918
A/Accession: I38972
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-826 <RES>
A/Cross-references: UNIPROT:Q16665; EMBL:U22431; NID:9881345; PIDN:AACG0152.1; PID:98813
A/Note: parts of this sequence were confirmed by peptide sequencing
R/Hogenssch, J.B.; Chan, W.K.; Carver, L.A.; Bradford, C.A. submitted to the EMBL data library, June 1995
A/Reference number: H00692
A/Accession: G01875
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: mRNA
A/Residues: 1-826 <HOG>
A/Cross-references: EMBL:U29165; NID:G1144012; PIDN:AAC51210.1; PID:G1144013
C/Genetics:
A/Gene: GDB:HIF1A
A/Cross-references: GDB:512229
A/Map position: 14q21-14q24
C/Keywords: heterodimer

Query Match 30.3%; Score 495; DB 2; Length 826;

Best Local Similarity 46.2%; Pred. No. 4.1e-35;
Matches 114; Conservative 26; Mismatches 59; Indels 48; Gaps 6;
QY 11 NTELRKERSDARSRQETEVLYQALATLPFARGVSAHLDKASIMRLTISTLRMRRLC 70
DB 14 SSRRREKSRDARSRKSESEVFYELAHQLPLPHVSSHLDDASVWRLTISTLRMRRLC 73
QY 71 AAGCKRGATGRLPEPGGFRHGTNRGRHGLPVGKCOQAPGQSVLDC 120
DB 74 DAG-----DLIDIDDKAAQMKNCYKALDGFVWVLTLD 105
QY 121 SSLIHT-----NPTPG-TNFSLELIGHSIFDFIHPDOELDALTPRPNLSSKYLEAFT 174
DB 106 DGMITISDNVKNYMGILTQF--ELTGHSVDFPTHPDHEHEMLTHRNGLYKCKEQNT 163
QY 175 ERHFSLRMSTLTSRGRITNLKAAITKYLHCSGMRAKYKRPNOTSPAGSPRSEPPLOCLV 234
DB 164 GRFFPLRMKCTLTSRGRITNLKAAITKYLHCSGMRAKYKRPNOTSPAGSPRSEPPLOCLV 221
QY 235 LVICEAIP 241
DB 222 LVICEPIP 228

RESULT 4
JC4837
hypoxia-inducible factor 1 alpha - mouse
C/Species: Mus musculus (house mouse)
C/Date: 15-Aug-1996 #sequence_revision 15-Oct-1996 #text_change 09-Jul-2004
C/Accession: JC4837
R/Wenger, R.H.; Rolfs, A.; Marti, H.H.; Guenet, J.L.; Gaasmann, M. Biochem. Biophys. Res. Commun. 223, 54-59, 1996
A/Title: Nucleotide sequence, chromosomal assignment and mRNA expression of mouse hypoxi
A/Reference number: JC4837; MUID:96254028; PMID:8660378
A/Accession: JC4837
A/Molecule type: mRNA
A/Residues: 1-810 <MEN>
A/Cross-references: UNIPROT:Q61221; EMBL:X05580; NID:G1430864; PIDN:CAA64833.1; PID:G437
A/Comment: This factor is involved in the oxygen-regulated transcription of several genes
C/Genetics: hif-1alpha
A/Gene: Hif1alpha
A/Map position: 12
A/Keywords: transcription factor
F/5-58/Region: helix-loop-helix #status predicted

Query Match 30.1%; Score 492; DB 2; Length 810;
Best Local Similarity 45.0%; Pred. No. 7.3e-35;
Matches 112; Conservative 26; Mismatches 59; Indels 52; Gaps 5;

QY 11 NTELRKERSDARSRQETEVLYQALATLPFARGVSAHLDKASIMRLTISTLRMRRLC 70
DB 2 SSRRREKSRDARSRKSESEVFYELAHQLPLPHVSSHLDDASVWRLTISTLRMRRLC 61
QY 71 AAGCKRGATG-----LPEPGGFRHGTNRGRHGLPVGKCOQAPG 113
DB 62 DAGGLDSEDEMKQMKCFYKALDGFVWVLTDDG----- 95
QY 114 PGQVDCSSSLIHNPTPG-TNFSLELIGHSIFDFIHPDOELDALTPRPNLSSKYLEA 172
DB 96 ----DMVYISDNVKNYMGILTQF--ELTGHSVDFPTHPDHEHEMLTHRNGVVRGKEL 149
QY 173 PTERHFSLRMSTLTSRGRITNLKAAITKYLHCSGMRAKYKRPNOTSPAGSPRSEPPLOCLV 232
DB 150 NTGRSFFLRMKCTLTSRGRITNLKAAITKYLHCSGMRAKYKRPNOTSPAGSPRSEPPLOCLV 207
QY 233 LVICEAIP 241
DB 208 LVICEPIP 216

RESULT 5
JC5809
hypoxia-inducible factor 1 alpha - rat

```

      127 MLQFVPSLRDNDIKQDIETAE-----DQGEVKEKLEVTGEDWINGAEAREL 174
      Db

```

```
QY      125 IHNPTGNTFSLLEIGHSTFDPIHPCDQESLQDALTPRPNISSKKL-EAPLERHFSLRMK 183
          :|||:::|||::|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      116 -----QVELTGNISLYEYIHPSDHDDEMTAVITAHQPLHHLLLOEYEIERSFFLRMK 165
```

Db 116 -----QVELTGSNIYEHPSDHDENTAVLTAHQPLHHLLQBYEIERSFFLRMK

Qy 184 STLSRGRTLNKKAATWKVHCSGMR 210
| : : : : :
Db 166 CVLAK--RNAGLTCSGYKVIHCSGYLK 190

RESULT 8

A29945
neurogenesis regulatory protein - fruit fly (Drosophila melanogaster) (fragment)
N:Alternate names: single-minded gene protein
C/Species: Drosophila melanogaster
C/Date: 15-Dec-1998 #sequence_revision 15-Dec-1998 #text_change 09-Jul-2004
C/Accession: A29945
R/Crews: S.T.; Thomas, J.B.; Goodman, C.S.
Cell 52, 143-151, 1988
A/Title: The Drosophila single-minded gene encodes a nuclear protein with sequence simi
A/Reference number: A29945; MUID:88151023; PMID:3345560
A/Accession: A29945
A/Molecule type: mRNA
A/Residues: 1-655 <CR>
A/Cross-references: UNIPROT:P05709; GB:M19020; NID:g158464; PID:g158465
C/Genetics:
A/Gene: sim
C/Keywords: DNA binding; transcription regulation

Query Match 12.3%; Score 202; DB 2; Length 655;
Best Local Similarity 28.5%; Pred. No. 1.2e-09;
Matches 73; Conservative 35; Mismatches 78; Indels 70; Gaps 10;

Qy 36 QLAHTLPFAGVSAHLDKASIMRLTISYLRMHRCAGAGKRGATGRLPEPGGPFHGT 95
| : : : : :
Db 4 ELAKLPLPAATISQLDKASVIRLTITSLKRRQVPDGG--LGEAWG---SSPAMQRGAT 57
Qy 96 HRR-GRHGL-----PVGKCOQAPGQSVYDLCSSSLINHPRTGTFSLIGHSI 143
| : : : : :
Db 58 IKELGSHLQTLDFGFVPAADGKIMYISERASVHLGSS-----QVELTNSI 105
Qy 144 PDLHPDQDELQALPRPPLSKKUL-----EAPL-----ERH 177
| : : : : :
Db 106 FEYIHNVDDDEMNILSLHPINOHPLAQHTTPIGSPGVQHPSPAYDHRGSHTEIEKT 165
Qy 178 FSLMKSTLSRGRTLNKKAATWKVHCSGMRVYKPPAQGSPGSRSEPPQLQCLVLC 237
| : : : : :
Db 166 FFLMKCVLAK--RNAGLTISGFVYHCSGYLKARITPDRGSGGS-----LIONLGLVA 218
Qy 238 -----EALPOLPFH 246
| : : : :
Db 219 VGHSLPSSATIEIKLH 234

RESULT 9

T21943
hypothetical protein F38A6.3a - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T21943; T26899
R/Mortimore, B.
submitted to the EMBL Data Library, March 1997
A/Reference number: Z19491
A/Accession: T21943
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-823 <WIL>
A/Cross-references: UNIPROT:O45486; EMBL:Z92833; PIDN:CAB07380.1; GSPDB:GN00023; CESP:F3
A/Experimental source: clone F38A6
R/Ainscough, R.
submitted to the EMBL Data Library, June 1998
A/Reference number: Z20282
A/Accession: T26899
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-823 <WIL>
A/Cross-references: EMBL:AL023842; PIDN:CAA19520.1; GSPDB:GN00023; CESP:F38A6.3a

A/Experimental source: clone Y44A6D
C/Genetics:
A/Gene: CESP:F38A6.3a
A/Map position: 5
A/Introns: 19/2; 75/1; 111/2; 175/1; 226/3; 331/3; 518/1; 627/2; 711/1; 768/3

Query Match 10.4%; Score 169.5; DB 2; Length 823;
Best Local Similarity 26.3%; Pred. No. 1.1e-06;
Matches 62; Conservative 35; Mismatches 90; Indels 49; Gaps 10;

Qy 9 RSNTELRKESRRAASRSQETEVLYOAHLPFA-RGVSAHLDKASIMRLTISYLRMH 67
| : : : : :
Db 110 KRMRERRRSTRRAADRSKESDIPDLKMCYPIYBEQTVTHLDIALLRVAATTCRLR 169
Qy 68 RLCA---AGGKRGATGRLPEPGGFRHGTNRGRHGLPVGKCOQAPGQSVYDLCSSTL 124
| : : : : :
Db 170 KTAGNVLENLNDNEITNEWTED-----TIAECLDG-FVMIVDSDSL 212
Qy 125 INHPRTGTFSL---ELIGHSTFDHPDQDELQALPRPPLSKKLEAP-----T 174
| : : : : :
Db 213 YVTESVAMYGLTQTLTGRLALDFLHPSDYDEF-----DKQSKMLHKPRGEDTDTT 264
Qy 175 ERHFSLRKSTLSRGRTLNKKAATWK---VLH-----CSGMRAYKPPAQGSPAG 222
| : : : : :
Db 265 GIMVLRMKTVISPRGRCLNLSALYKSVFLVHSVSTGSHVSPWG--GITIPAG 318

RESULT 10

T21944
hypothetical protein F38A6.3b - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T21944; T26900
R/Mortimore, B.
submitted to the EMBL Data Library, March 1997
A/Reference number: Z19491
A/Accession: T21944
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-825 <WIL>
A/Cross-references: UNIPROT:Q9TVN0; EMBL:Z92833; PIDN:CAB07381.1; GSPDB:GN00023; CESP:F3
A/Experimental source: clone F38A6
R/Ainscough, R.
submitted to the EMBL Data Library, June 1998
A/Reference number: Z20282
A/Accession: T26900
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-825 <WIL>
A/Cross-references: EMBL:AL023842; PIDN:CAA19521.1; GSPDB:GN00023; CESP:F38A6.3b
A/Experimental source: clone Y44A6D
A/Genetics:
A/Gene: CESP:F38A6.3b
A/Map position: 5
A/Introns: 19/2; 75/1; 111/2; 175/1; 226/3; 331/3; 518/1; 627/2; 713/1; 770/3

Query Match 10.4%; Score 169.5; DB 2; Length 825;
Best Local Similarity 26.3%; Pred. No. 1.1e-06;
Matches 62; Conservative 35; Mismatches 90; Indels 49; Gaps 10;

Qy 9 RSNTELRKESRRAASRSQETEVLYOAHLPFA-RGVSAHLDKASIMRLTISYLRMH 67
| : : : : :
Db 110 KRMRERRRSTRRAADRSKESDIPDLKMCYPIYBEQTVTHLDIALLRVAATTCRLR 169
Qy 68 RLCA---AGGKRGATGRLPEPGGFRHGTNRGRHGLPVGKCOQAPGQSVYDLCSSTL 124
| : : : : :
Db 170 KTAGNVLENLNDNEITNEWTED-----TIAECLDG-FVMIVDSDSL 212
Qy 125 INHPRTGTFSL---ELIGHSTFDHPDQDELQALPRPPLSKKLEAP-----T 174
| : : : : :
Db 213 YVTESVAMYGLTQTLTGRLALDFLHPSDYDEF-----DKQSKMLHKPRGEDTDTT 264
Qy 175 ERHFSLRKSTLSRGRTLNKKAATWK---VLH-----CSGMRAYKPPAQGSPAG 222

Db 265 GINNVLRKTVISPRGCLINKLSALYKSVFLVSHSKYSTGSHVSPMQ--GITTPAG 318

RESULT 11

JC7635

aryl hydrocarbon receptor nuclear translocator 1 - chicken

N/Alternate names: hypoxia-inducible factor 1 beta

C/Species: Gallus gallus (chicken)

C/Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 09-Jul-2004

C/Accession: JC7635

R/Catron, T.; Mendiola, M.A.; Smith, S.M.; Born, J.; Walker, M.K.

Biochem. Biophys. Res. Commun. 282, 602-607, 2001

A/Title: Hypoxia regulates avian cardiac Arnt and HIF-1alpha mRNA expression.

A/Reference number: JC7635; PMID:21294777; PMID:11401503

A/Contents: Embryo

A/Accession: JC7635

A/Molecule type: mRNA

A/Residues: 1-805 <CAT>

A/Cross-references: UNIPROT:Q985N3

C/Comment: This protein, known as hypoxia-inducible factor 1 beta, a member of the basic pones to hormonal and environmental stimuli. It is also a common dimer partner for trar

C/Genetics:

A/Gene: arntl

C/Keywords: transcription factor

Query Match 8.7%; Score 142; DB 2; Length 805;

Best Local Similarity 20.5%; Pred. No. 0.00027;

Matches 61; Conservative 43; Mismatches 95; Indels 98; Gaps 8;

Qy 14 LRKESRDAAARRSQETEVLYOLAHITLPPARGVSAHLDAKASIMRLTISYLRMRLCAAG 73

Db 89 LARENHSEIERRRRNKMTAYITELSDWVPTCSALARKPKDLTILRMVSHM----- 139

Qy 74 GKRGRATGRLLPBGPGFGRHGRHGLPVKCCQAQPGQSVLDCSSSLI---HNPT 129

Db 140 -KSLRGTGNTSTGSIKPSFLTDQLKHLI---LEADGFLIVSCETGRVYVSDSVT 194

Qy 130 PGTNF-SLELIGHSIFPFIHPCDOELQDALTRPN-----LSKKLEAPTERHF 178

Db 195 PVLNQPSWFGSTLYDQVHPDDVTKLRQLSTSENLTEGTXPWCLSNKPAPPS-- 252

Qy 179 SLRKSTLTLSRGRITLNLKATWK----- 201

Db 253 -----ASKGRILDLKTGTVKKEGQSMKCMGSRRSFICMRGNSVDVAVNRLS 304

Qy 202 -----VHCSGMRAYKPPAQSPSPGSP-RSEPPQLCVLI 236

Db 305 FMNRNCRNGLGATYDGEHPHYVAVHCTGYIKAMPAGVSLPDDPDAGGSKFCLVAI 361

RESULT 12

JC7633

aryl hydrocarbon nuclear translocator ARNT2-like factor, ARNT2X - zebra fish

C/Species: Brachydanio rerio (zebra fish)

C/Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 09-Jul-2004

C/Accession: JC7633

R/Hu, H.T.; Wang, W.D.; Hu, C.H.

Biochem. Biophys. Res. Commun. 282, 487-492, 2001

A/Title: Ectopic expression of negative ARNT2 factor disrupts fish development.

A/Reference number: JC7633; PMID:21294759; PMID:11401485

A/Contents: Heart

A/Accession: JC7633

A/Molecule type: mRNA

A/Residues: 1-392 <HSU>

A/Cross-references: UNIPROT:Q985K3; GB:AY007992

C/Comment: This factor, a basic helix-loop-helix PAS (bHLH-PAS) factor, heterodimerizes

C/Genetics:

A/Gene: arnt2X

Query Match 8.5%; Score 138.5; DB 2; Length 392;

Best Local Similarity 20.8%; Pred. No. 0.00023;

Matches 55; Conservative 48; Mismatches 103; Indels 59; Gaps 8;

Qy 16 KEKSRDAARRSQETEVLYOLAHITLPPARGVSAHLDAKASIMRLTISYLRMRLCAAGK 75

Db 47 RENHSEIERRRRNKMTQYTIELSDWVPTCSALARKPKDLTILRMVSHM-----K 96

Qy 76 RGRATGRLLPBGPGFGRHGRHGLPVKCCQAQPGQSVLDCSSSLI---HNPTPG 131

Db 97 SMRGTGNTSTGAYKPSFLTEQLKHLI---LEADGFLFVVAETGRVYVSDSVTPV 152

Qy 132 TNF-SLELIGHSIFPFIHPCDOELQDALTRPNLSKKL----- 170

Db 153 LNHQSWFGSTLYDQVHPDDVTKLRQLSTSENLTEGTXPWCLSNKPAPPS-- 212

Qy 171 -----EAPTERHFSLRMSTLTSTGRGRT-----LNLKVAATWXYLCSGHRAY 212

Db 213 MGRSRSFICMRGNSALD-HISLNRSSMRKRYRNLGSPSKGEAQYVHCTGYIKAW 271

Qy 213 KPFAQTSP-AGSPREPPQLCVLI 236

Db 272 PPAGMTIPDDPDTEAGQTSKYCLVAI 296

RESULT 13

A56241

aryl hydrocarbon receptor nuclear translocator protein - mouse

C/Species: Mus musculus (house mouse)

C/Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 09-Jul-2004

C/Accession: A56241

R/Reisz-Porszasz, S.; Prober, M.R.; Fukunaga, B.N.; Hankinson, O.

Mol. Cell. Biol. 14, 6075-6086, 1994

A/Title: Identification of functional domains of the aryl hydrocarbon receptor nuclear t

A/Reference number: A56241; PMID:94344118; PMID:8065341

A/Accession: A56241

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-791 <REI>

A/Cross-references: UNIPROT:P53762; GB:U10325; NID:9555687; PID:AAA56717.1; PID:9555688

A/Note: authors translated the codon TTT for residue 375 as Ser, TGT for residue 380 as s

A/Note: authors failed to translate CAG for residue 507 as Gln

C/Genetics:

A/Gene: ARNT

Query Match 8.5%; Score 138.5; DB 2; Length 791;

Best Local Similarity 21.8%; Pred. No. 0.00053;

Matches 58; Conservative 50; Mismatches 101; Indels 57; Gaps 9;

Qy 14 LRKESRDAAARRSQETEVLYOLAHITLPPARGVSAHLDAKASIMRLTISYLRMRLCAAG 73

Db 89 LARENHSEIERRRRNKMTAYITELSDWVPTCSALARKPKDLTILRMVSHM----- 139

Qy 74 GKRGRATGRLLPBGPGFGRHGRHGLPVKCCQAQPGQSVLDCSSSLI---HNPT 129

Db 140 -KSLRGTGNTSTGSIKPSFLTDQLKHLI---LEADGFLIVSCETGRVYVSDSVT 194

Qy 130 PGTNF-SLELIGHSIFPFIHPCDOELQDALTRPN-----LSKKLEAPT----- 174

Db 195 PVLNQPSWFGSTLYDQVHPDDVTKLRQLSTSENLTEGTXPWCLSNKPAPPS-- 254

Qy 175 -----EHSLSLRMSTLTST-----RGRTLNLKAT-----WVYLHCSGMR 211

Db 255 MCMGSRRSFICMRGNSVDVPMNRLSFLNRNCRNGLSVXGEBHPFVAVHCTGYIKA 314

Qy 212 YKPPAQTSPPAGSPRS--BPPLQCVLI 236

Db 315 WPRAGVSLPDDPDTEAGGSKFCLVAI 340

RESULT 14

A55448

Ah receptor nuclear translocator - mouse

C/Species: Mus musculus (house mouse)

C/Date: 10-Feb-1995 #sequence_revision 10-Feb-1995 #text_change 09-Jul-2004

C/Accession: A55448

R, Li, H.; Dong, L.; Whitlock Jr., J.P.
J. Biol. Chem. 269, 28098-28105, 1994
A>Title: Transcriptional activation function of the mouse Ah receptor nuclear translocator
A/Reference number: A55448; MUID:95050586; PMID:7961746
A/Accession: A55448
A>Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-776 <LIA>
A/Cross-references: UNIPROT:Q921F3; GB:U14333

Query Match 8.3%; Score 136.5; DB 2; Length 776;
Best Local Similarity 21.2%; Pred. No. 0.00077;
Matches 58; Conservative 50; Mismatches 109; Indels 57; Gaps 9;

QY 6 QRVASNTLREKRSRDARSRSQETEVLYQLAHTLPFARGVSAHLDKASIMRLTISYLR 65
DB 66 QMCDKERPARENSEIERRRRNMTAVITELSDMVPFCSALARKPKDLTLRMAVSHM- 124
QY 66 MHRICAAAGKRGATGRLLPBGPGFRHGTNRGRHGLPVGKCOQAPGQSVDLCSLSLI 125
DB 125 -----KSLRGCTGNTSTGSKYKPSFLTDQELKHLI-----LEADGFLFIVSCETGRV 171
QY 126 ----HNPTPGTNF-SLELIGHSTFDLHPCDQELDALTFRPV-LSKKKLEAPT----- 174
DB 172 VYYSDSVTPVLNQPSQSWFGSTLYDQVHPDQVDRKLEQLSTSENALTGRLDLKTGYKK 231
QY 175 -----ERHFSLRMKSTLTS-----RGRTLNLKAAAT-----WKVL 203
DB 232 EGOQSSMRMCMGSRSPFCMRRCCTSSVDVPVSMRLSFLNRCNGLGSKVKGPHFVVV 291
QY 204 HCSGMRAYKPPACTSPAGSPRS-EPPLOCLVLI 236
DB 292 HCTGYIKAMPAGVSLPDDPDEAGQSKFCLVAI 325

RESULT 15

I59550
aryl hydrocarbon receptor nuclear translocator Arnt [imported] - human
C/Species: Homo sapiens (man)
C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C/Accession: I59550
R/Hoffman, E.C.; Reyes, H.; Chu, F.F.; Sander, F.; Conley, L.H.; Brooks, B.A.; Hankinson
Science 252, 954-958, 1991
A>Title: Cloning of a factor required for activity of the Ah (dioxin) receptor.
A/Reference number: I59550; MUID:91240280; PMID:1852076
A/Accession: I59550
A>Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-789 <RBS>
A/Cross-references: UNIPROT:P27540; GB:M69238; NID:G179003; PIDN:AAA51777.1; PID:G179004
C/Genetics:
A/Gene: GDB:ARNT
A/Cross-references: GDB:119701; OMIM:126110
A/Map position: 1q21-1q21

Query Match 8.3%; Score 136.5; DB 2; Length 789;
Best Local Similarity 21.8%; Pred. No. 0.00079;
Matches 58; Conservative 50; Mismatches 101; Indels 57; Gaps 9;

QY 14 LRKESDAARSRSQETEVLYQLAHTLPFARGVSAHLDKASIMRLTISYLRMRLCAAG 73
DB 89 LARENHSEIERRRNMTAVITELSDMVPFCSALARKPKDLTLRMAVSHM----- 139
QY 74 GKRGRATGRLLPBGPGFRHGTNRGRHGLPVGKCOQAPGQSVDLCSLSLI-----HNPT 129
DB 140 -KSLRGCTGNTSTGSKYKPSFLTDQELKHLI-----LEADGFLFIVSCETGRVYVSDSVT 194
QY 130 PGTNF-SLELIGHSTFDLHPCDQELDALTFRPV-LSKKKLEAPT----- 174
DB 195 PVLNQPSQSWFGSTLYDQVHPDQVDRKLEQLSTSENALTGRLDLKTGYKKEGOQSSMR 254
QY 175 -----ERHFSLRMKSTLTS-----RGRTLNLKAAAT-----WKVLHCSGMR 211

DB 255 MCKGRRSPFCMRRCSSSVDPVSVNRLSFVRRNRCNGLGSKYKDGEPHFVWHCTGYKA 314
QY 212 YKPPACTSPAGSPRS-EPPLOCLVLI 236
DB 315 WPPAGVSLPDDPDEAGQSKFCLVAI 340

Search completed: June 15, 2005, 16:00:54
Job time: 42 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 15:37:58 ; Search time 173 Seconds
(without alignments)
908.719 Million cell updates/sec

Title: US-09-896-791B-3
Perfect score: 1636
Sequence: 1 MALGLQVRNTELRKESR.....TESSLPSWVLMALNRKNCPG 307

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1636	100.0	307	Q8VHR1	Q8vhr1 mus musculus
2	855.5	52.3	630	Q9QX54	Q9qxs4 mus musculus
3	855.5	52.3	662	Q9Z2I5	Q9z2i5 mus musculus
4	835	51.0	632	Q8WXA1	Q8wx1 homo sapien
5	835	51.0	669	Q6KX72	Q6kx72 homo sapien
6	831	50.8	662	Q96K34	Q96k34 homo sapien
7	820.5	50.2	662	Q9JH52	Q9jh52 ratius norv
8	802	49.0	667	Q9Y2N7	Q9y2n7 homo sapien
9	573	35.0	648	Q9HA12	Q9ha12 homo sapien
10	518.5	31.7	199	Q9HAMS	Q9hams homo sapien
11	503	30.7	811	HIFA_CHICK	O9yib9 gallus gall
12	499	30.5	823	HIFA_BOVIN	O9xtab bos taurus
13	499	30.5	823	HIFA_RAT	O6iv47 bos mutus g
14	497	30.4	825	HIFA_RAT	O35800 ratius norv
15	495	30.3	489	Q6EMT3	O6emi3 xenopus lae
16	495	30.3	826	HIFA_HUMAN	O6i665 homo sapien
17	495	30.3	836	HIFA_MOUSE	O6i221 mus musculus
18	493	30.1	802	Q6P154	Q6p154 xenopus lae
19	492.5	30.1	643	Q6STN6	Q6stn6 ctenopharyn
20	492	30.1	819	Q7YSES	Q7yses cycloclagus
21	490.5	30.0	777	Q6EH14	Q6eh14 brachydanio
22	489.5	29.9	774	Q6STN7	Q6stn7 ctenopharyn
23	488	29.8	874	PAS1_MOUSE	P97481 mus musculus
24	488	29.8	874	Q6PBU2	Q6pbu2 mus musculus
25	484	29.6	821	Q64FS4	Q64fs4 spermophilu
26	483.5	29.6	533	Q6P133	Q6p133 brachydanio
27	482	29.5	874	PAS1_RAT	O9jhai ratius norv
28	481.5	29.4	626	Q6EGR9	Q6egr9 brachydanio
29	480	29.3	786	Q6SLI1	O6sli1 canis famil
30	479	29.3	835	Q66W2	Q66w2 ctenopharyn
31	478	29.2	824	Q6H8T3	Q6h8t3 spalax juda

32	474.5	29.0	873	Q8QGM4	Q8qgm4 fundulus he
33	471	28.8	862	Q6GL61	Q6gl61 xenopus tro
34	471	28.8	867	Q9W7C6	Q9w7c6 gallus gall
35	471	28.8	870	Q9PTB3	Q9ptb3 coturnix co
36	468	28.6	870	Q9XTA4	Q9xta4 bos taurus
37	467	28.5	870	PAS1_HUMAN	Q99814 homo sapien
38	464	28.4	862	Q6GP97	Q6gp97 xenopus lae
39	462	28.2	859	Q6GQ12	Q6gq12 xenopus lae
40	450	27.5	766	HIFA_ONCMY	Q98sw2 oncorhynch
41	425	26.0	237	Q6H7Z9	Q6h7z9 homo sapien
42	389	23.8	189	Q6RYC8	Q6ryc8 ovis aries
43	376	23.0	189	Q6RYC9	Q6ryc9 sus scrofa
44	360	22.0	65	Q8R4D6	Q8r4d6 mus musculus
45	323.5	19.8	571	Q7T2E4	Q7t2e4 brachydanio

ALIGNMENTS

RESULT 1					
ID	Q8VHR1	PRELIMINARY;	PRT;	307 AA.	
AC	Q8VHR1				
DT	01-MAR-2002 (TREMBLrel. 20, Created)				
DT	01-MAR-2002 (TREMBLrel. 20, Last sequence update)				
DE	01-MAR-2004 (TREMBLrel. 26, Last annotation update)				
DE	Inhibitory PAS domain protein.				
GN	Name:Hifa3, Synonyms:Idas;				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=C57BL/6J;				
RX	PubMed=11734856; DOI=10.1038/35107085;				
RA	Makino Y., Cao R., Svensson K., Bertilsson G., Asman M., Tanaka H.,				
RA	Cao Y., Berkenstam A., Poellinger L.;				
RT	"Inhibitory PAS domain protein is a negative regulator of hypoxia-				
RT	inducible gene expression."				
RL	Nature 414:550-554(2001).				
CC	-1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.				
DR	EMBL; AF416641; ALJ39015.1; -.				
DR	MCD; MGI:1859778; Hf3a.				
DR	GO; GO:0005634; C:nucleus; IC.				
DR	GO; GO:0003700; P:transcription factor activity; IPI.				
DR	GO; GO:0001666; P:response to hypoxia; IDA.				
DR	Pfam; PF00010; HLM; 1.				
DR	SMART; SM00353; HLM; 1.				
SQ	SEQUENCE 307 AA; 33990 MW; D89D3B225C9C3967 CRC64;				
Query Match					
Best Local Similarity 100.0%; Score 1636; DB 2; Length 307;					
Matches 307; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	MALGLQVRNTELRKESRDAARSRSOSTEVLVYOLAHLPAPRGVSAHLDKASIRLT	60		
DB	1	MALGLQVRNTELRKESRDAARSRSOSTEVLVYOLAHLPAPRGVSAHLDKASIRLT	60		
QY	61	ISYLRMRLLCAAGKGRATGRLLPEGPGRFRGTHRGHGLPVGKCOQAPGOSVDLC	120		
DB	61	ISYLRMRLLCAAGKGRATGRLLPEGPGRFRGTHRGHGLPVGKCOQAPGOSVDLC	120		
QY	121	SSSLINHPGTNFSLELIGHSIFDFIHPCDQBELDALTPRPVLSKKKEAPTERFSL	180		
DB	121	SSSLINHPGTNFSLELIGHSIFDFIHPCDQBELDALTPRPVLSKKKEAPTERFSL	180		
QY	181	RMSSTLSRGRTNLKATKVLHCSGMRAYVPAQTSFAGSPRSPPLQCVLICEAI	240		
DB	181	RMSSTLSRGRTNLKATKVLHCSGMRAYVPAQTSFAGSPRSPPLQCVLICEAI	240		
QY	241	POLPFHDGATLGPQEKTPISTLFTPLMKALLCLVKKMPVVOLOGKTESSLPSWVLMAL	300		

Db 241 POLPFHGGATGIPQEKPISTLPPLMKALCLVKEWPVQVLOGKTESLPSWLMAL 300
 QY 301 NRKNCPG 307
 Db 301 NRKNCPG 307

RESULT 2

090X54 PRELIMINARY; PRT; 630 AA.

AC 090X54; (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 26, Last annotation update)
 DE Hypoxia-inducible factor 3 alpha (Fragment).
 GN Name=Hif3a;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Gu Y.-Z., Moran S.M., Hogenesch J.B., Wartman L., Bradfield C.A.;
 RT "Cloning and Characterization of a Third Hypoxia Inducible Factor,
 RT HIF3-alpha."
 RL J. Biol. Chem. 0:0-0(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99054547; PubMed=9840812;
 RA Gu Y.-Z., Moran S.M., Hogenesch J.B., Wartman L., Bradfield C.A.;
 RT "Molecular characterization and chromosomal localization of a third
 RT alpha-class hypoxia inducible factor subunit, HIF3alpha."
 RL Gene Expr. 7:205-213(1998).
 CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
 DR EMBL: AF079153; AAF21782.1; JOINED.
 DR EMBL: AF079140; AAF21782.1; JOINED.
 DR EMBL: AF079141; AAF21782.1; JOINED.
 DR EMBL: AF079143; AAF21782.1; JOINED.
 DR EMBL: AF079145; AAF21782.1; JOINED.
 DR EMBL: AF079147; AAF21782.1; JOINED.
 DR EMBL: AF079149; AAF21782.1; JOINED.
 DR EMBL: AF079151; AAF21782.1; JOINED.
 DR EMBL: AF079152; AAF21782.1; JOINED.
 DR EMBL: AF079150; AAF21782.1; JOINED.
 DR EMBL: AF079148; AAF21782.1; JOINED.
 DR EMBL: AF079146; AAF21782.1; JOINED.
 DR EMBL: AF079144; AAF21782.1; JOINED.
 DR EMBL: AF079142; AAF21782.1; JOINED.
 DR HSSP: Q99814; 1P97.
 DR MGD; MGI:1859778; Hif3a.
 DR GO: GO:0005634; C:nucleus; IC.
 DR GO: GO:0003700; F:transcription factor activity; IPI.
 DR GO: GO:0001666; P:response to hypoxia; IDA.
 DR GO: GO:0006366; P:transcription from Pol II promoter; IPI.
 DR InterPro: IPR001092; HLH_basic.
 DR InterPro: IPR001067; NUC_translocat.
 DR InterPro: IPR001610; PAC.
 DR InterPro: IPR000014; PAS.
 DR Pfam; PF000010; HLH; 1.
 DR Pfam; PF00989; PAS; 1.
 DR PRINTS; PR00785; NCTRNSLOCATR.
 DR SMART; SM00353; HLH; 1.
 DR SMART; SM00086; PAC; 1.
 DR SMART; SM00091; PAS; 2.
 DR PROSITE; PSS0112; PAS; 2.
 FT NON_TER 630 630
 SQ SEQUENCE 630 AA; 69623 MW; 828EB2CB4B6D45B6 CRC64;

Query March 52.3%; Score 855.5; DB 2; Length 630;
 Best Local Similarity 68.4%; Pred. No. 1.8e-62;
 Matches 182; Conservative 5; Mismatches 24; Indels 55; Gaps 4;

QY 9 RSTTELRKESKRDPAASRSQETEVLYOLAHITPPARGVSAHLDDKASIMRLTISYLMHR 68
 Db 7 RSTTELRKESKRDPAASRSQETEVLYOLAHITPPARGVSAHLDDKASIMRLTISYLMHR 66
 QY 69 LCAAGGGRGATGRLPEGGGGRHGRGRGLPVGKQQAPEGQSVLDC----- 120
 Db 67 LCAAGGWN-----QVEKGGEPPDCTYKALGEF 94
 QY 121 -----SSSLINHPPTGTFNS-DELIGHSIFDFIHPDQDELQDALTPRNLSSKKL 170
 Db 95 VMLVLTAGDMAYLSENVSKLGLSQLELIGHSIFDFIHPDQDELQDALTPRNLSSKKL 154
 QY 171 EAPTEHFSLRKSTLTSGRTINLKAATWKVLHSGHNRAYRPPAQTSPAGSPRSEPL 230
 Db 155 EAPTEHFSLRKSTLTSGRTINLKAATWKVLHSGHNRAYRPPAQTSPAGSPRSEPL 214
 QY 231 QCVLICEATP-----QLPFGDGL 251
 Db 215 QCVLICEATPFPASLEPPLGRGAPL 240

RESULT 3

092Z15 PRELIMINARY; PRT; 662 AA.

AC 092Z15; (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-MAY-2004 (TREMBlrel. 26, Last annotation update)
 DE Hypoxia inducible factor three alpha.
 GN Name=Hif3a;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99054547; PubMed=9840812;
 RA Gu Y.-Z., Moran S.M., Hogenesch J.B., Wartman L., Bradfield C.A.;
 RT "Molecular characterization and chromosomal localization of a third
 RT alpha-class hypoxia inducible factor subunit, HIF3alpha."
 RL Gene Expr. 7:205-213(1998).
 CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
 DR EMBL: AF060194; AAC72734.1; -.
 DR HSSP: Q99814; 1P97.
 DR MGD; MGI:1859778; Hif3a.
 DR GO: GO:0005634; C:nucleus; IC.
 DR GO: GO:0003700; F:transcription factor activity; IPI.
 DR GO: GO:0001666; P:response to hypoxia; IDA.
 DR GO: GO:0006366; P:transcription from Pol II promoter; IPI.
 DR InterPro: IPR001092; HLH_basic.
 DR InterPro: IPR001067; NUC_translocat.
 DR InterPro: IPR001610; PAC.
 DR InterPro: IPR000014; PAS.
 DR Pfam; PF000010; HLH; 1.
 DR Pfam; PF00989; PAS; 1.
 DR PRINTS; PR00785; NCTRNSLOCATR.
 DR SMART; SM00353; HLH; 1.
 DR SMART; SM00086; PAC; 1.
 DR SMART; SM00091; PAS; 2.
 DR PROSITE; PSS0112; PAS; 2.
 SQ SEQUENCE 662 AA; 73012 MW; 58740A1B6993D3B5 CRC64;

Query Match 52.3%; Score 855.5; DB 2; Length 662;
 Best Local Similarity 68.4%; Pred. No. 2e-62;
 Matches 182; Conservative 5; Mismatches 24; Indels 55; Gaps 4;

QY 9 RSTTELRKESKRDPAASRSQETEVLYOLAHITPPARGVSAHLDDKASIMRLTISYLMHR 68
 Db 7 RSTTELRKESKRDPAASRSQETEVLYOLAHITPPARGVSAHLDDKASIMRLTISYLMHR 66
 QY 69 LCAAGGGRGATGRLPEGGGGRHGRGRGLPVGKQQAPEGQSVLDC----- 120
 Db 67 LCAAGGWN-----QVEKGGEPPDCTYKALGEF 94

Qy 121 -----SSSLINHPPTGNTFS-LELIGHSIPDFIHPDQEBLQDALTPRPNLSSKXL 170
 Db 95 VVWVLTAEQDMVYLSNVSKHGLSGLLELIGHSIPFIFHPDQEBLQDALTPRPNLSSKXL 154
 Qy 171 EAPFERHSLRMKSTLTSGRTLNKAAATWVYLHCSGMRAYKPPAQTSPPAGSPSEPP 230
 Db 155 EAPFERHSLRMKSTLTSGRTLNKAAATWVYLHCSGMRAYKPPAQTSPPAGSPSEPP 214
 Qy 231 QCLVLTCEAIP-----OLPFDGATL 251
 Db 215 QCLVLTCEAIPHPASLEPPLGRGATL 240

RESULT 4

08WKX1 PRELIMINARY; PRT: 632 AA.
 AC 08WKX1
 DT 01-MAR-2002 (TRENBLrel. 20, Created)
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
 DE Inhibitory PAS domain protein.
 GN Name:IPAS;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Cheng J.Q.;
 RL Submitted (DRC-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
 DR EMBL; AF43492; AAL69947.1; -.
 DR HSSP; Q16655; 1IQB
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0004871; F:signal transducer activity; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR GO; GO:0007165; P:signal transduction; IEA.
 DR Pfam; PF00010; HLH; 1.
 DR Pfam; PF00989; PAS; 1.
 DR PRINTS; PR00785; NCTRNLOCATR.
 DR SMART; SM00353; HLH; 1.
 DR SMART; SM0086; PAC; 1.
 DR SMART; SM00091; PAS; 2.
 DR PROSITE; PS50112; PAS; 2.
 SQ SEQUENCE 632 AA; 68963 MW; 9665B0AF3998F8EF CRC64;

Query Match 51.0%; Score 835; DB 2; Length 632;
 Best Local Similarity 68.6%; Pred. No. 9.3e-61;
 Matches 179; Conservative 11; Mismatches 29; Indels 42; Gaps 5;
 Qy 1 MALGLQVRNSTEIRKESRDARSRSQETEVLYQLAHTLPFARGVSAHLDKASIMRLT 60
 Db 1 MALGLQVRNSTEIRKESRDARSRSQETEVLYQLAHTLPFARGVSAHLDKASIMRLT 60
 Qy 61 ISYLRMRRLCAAGGKRGRATGRLLPEBPGFRHOTHGRGHGRLPVGKQ----- 109
 Db 61 ISYLRMRRLCAAG-----EKNVGAAGEPDLACYLKALGFWVL 100
 Qy 110 QAPQSVVDLSSSLINHPPTGNTFS-LELIGHSIPDFIHPDQEBLQDALTPRPNLSSK 168
 Db 101 TAEGDMV-----LSENVSKHGLSGLLELIGHSIPFIFHPDQEBLQDALTPQOTLSR 154
 Qy 169 KLEAPTERHSLRMKSTLTSGRTLNKAAATWVYLHCSGMRAYKPPAQTSPPAGSPSEPP 228
 Db 155 KVEAPTERCFSLRMKSTLTSGRTLNKAAATWVYLHCSGMRAYKPPAQTSPPAGSPSEPP 214
 Qy 229 PLOGLVLTCEAIPOLPFHDA 249
 Db 215 PLOGLVLTCEAIP-----HPGS 231

RESULT 5

06KX72 PRELIMINARY; PRT: 669 AA.
 AC 06KX72
 DT 25-OCT-2004 (TRENBLrel. 28, Created)
 DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)
 DE HIF3A protein.
 GN Name:HIF3A;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;
 RX PubMed:12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Ditchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Muliyil S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Heintz B., Kettelman W., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Kravinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Maitra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;
 RA Director MGC Project;
 RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
 DR EMBL; BC080551; AAB80551.1; -.
 DR InterPro; IPR001092; HLH_basic.
 DR InterPro; IPR001067; NUC_translocat.
 DR InterPro; IPR000014; PAS.
 DR Pfam; PF00010; HLH; 1.
 DR PRINTS; PR00785; NCTRNLOCATR.
 DR SMART; SM00353; HLH; 1.
 DR SMART; SM00091; PAS; 2.
 DR PROSITE; PS50112; PAS; 2.
 SQ SEQUENCE 669 AA; 72460 MW; 7EF7362A1691AED CRC64;

Query Match 51.0%; Score 835; DB 2; Length 669;
 Best Local Similarity 68.6%; Pred. No. 1e-60;
 Matches 179; Conservative 11; Mismatches 29; Indels 42; Gaps 5;
 Qy 1 MALGLQVRNSTEIRKESRDARSRSQETEVLYQLAHTLPFARGVSAHLDKASIMRLT 60
 Db 1 MALGLQVRNSTEIRKESRDARSRSQETEVLYQLAHTLPFARGVSAHLDKASIMRLT 60
 Qy 61 ISYLRMRRLCAAGGKRGRATGRLLPEBPGFRHOTHGRGHGRLPVGKQ----- 109
 Db 61 ISYLRMRRLCAAG-----EKNVGAAGEPDLACYLKALGFWVL 100
 Qy 110 QAPQSVVDLSSSLINHPPTGNTFS-LELIGHSIPDFIHPDQEBLQDALTPRPNLSSK 168
 Db 101 TAEGDMV-----LSENVSKHGLSGLLELIGHSIPFIFHPDQEBLQDALTPQOTLSR 154
 Qy 169 KLEAPTERHSLRMKSTLTSGRTLNKAAATWVYLHCSGMRAYKPPAQTSPPAGSPSEPP 228
 Db 155 KVEAPTERCFSLRMKSTLTSGRTLNKAAATWVYLHCSGMRAYKPPAQTSPPAGSPSEPP 214

QY 229 PLOCVIVCEAIPQLPFDGA 249
 Db 215 PLOCVIVCEAIP----HPGS 231

RESULT 6

ID Q96K34 PRELIMINARY; PRT; 632 AA.

AC Q96K34
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Hypothetical protein FLJ14819.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary;
 RC PubMed=14702039; DOI=10.1038/ng1285;
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
 RA Sekine M., Ohayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
 RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K.,
 RA Murakami K., Yaeuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
 RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
 RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
 RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
 RA Tanai H., Kimata M., Watanabe M., Hirakawa S., Chiba Y., Ishida S.,
 RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
 RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
 RA Togiyia S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
 RA Musashino K., Yuuki H., Oshima A., Sasaki N., Aoyama S.,
 RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
 RA Motiwa S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujihara T.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hizo M., Ohmori Y.,
 RA Kaasabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Oktani R., Kawakami T., Noguchi S., Itoh T., Shigeta T., Senba T.,
 RA Matsunura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
 RA Togaishi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.,
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs.";
 RL Nat. Genet. 36:40-45 (2004).
 CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
 CC EMBL: AK027725; BAB55324.1; -
 DR HSBP; Q99814; 1P97.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0004871; F:signal transducer activity; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR GO; GO:0007165; P:signal transduction; IEA.
 DR InterPro; IPR001092; HLH_basic.
 DR InterPro; IPR001067; NUC_translocat.
 DR InterPro; IPR001610; PAC.
 DR InterPro; IPR000014; PAS.
 DR Pfam; PF00010; HLH; 1.
 DR Pfam; PF00989; PAS; 1.
 DR PRINTS; PR00785; NCTRNLOCATR.
 DR SMART; SM00353; HLH; 1.
 DR SMART; SM00086; PAC; 1.
 DR SMART; SM00091; PAS; 2.
 DR PROSITE; PSS0112; PAS; 2.
 SQ SEQUENCE 632 AA; 68933 MW; A19F1ED3D05E7A71 CRC64;

Query Match 50.8%; Score 831; DB 2; Length 632;
 Best Local Similarity 68.2%; Pred. No. 2e-60;
 Matches 178; Conservative 11; Mismatches 30; Indels 42; Gaps 5;

QY 1 MATGLQVRVRSNTLRKESGRDARSRSQETEVLYQLAHTLPARGVSAHLDKASITRLT 60
 Db 1 MATGLQVRVRSNTLRKESGRDARSRSQETEVLYQLAHTLPARGVSAHLDKASITRLT 60

QY 61 ISYLRMRRLCAAGKGRATGRLLPFGGGRFGRTHRGHGLPVGCKQ----- 109
 Db 61 ISYLRMRRLCAAG-----EMNQVAGGGRPLDACYLKALEGFVTVL 100

QY 110 QAPGQSVDCSSSLHNPTGTFNS-LELIGHSIFDPHPCQEEIQDALTPRPVLSKK 168
 Db 101 TAGGDMAV-----LSENVSKHLGLSQLELIGHSIFDPHPCQEEIQDALTPRQQLSRR 154

QY 169 KLEAPTRERHSLMRKSTLTSRGRTLVNKAATWVVLNCSGMRAYKPAQSPAGSPSEP 228
 Db 155 KVEAPTRERCSLMRKSTLTSRGRTLVNKAATWVVLNCSGMRAYKPAQSPAGSPSEP 214

QY 229 PLOCVIVCEAIPQLPFDGA 249
 Db 215 PLOCVIVCEAIP----HPGS 231

RESULT 7

ID Q9JH52 PRELIMINARY; PRT; 662 AA.

AC Q9JH52
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Hypoxia inducible factor 3 alpha.
 GN Name=Hif-3a;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_TaxID=10116;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RC MEDLINE=21134367; PubMed=11237857; DOI=10.1042/0264-6021.3540531;
 RA Kleczman T., Cornesse Y., Brechtel K., Modaresi S., Jungermann K.,
 RT "Perivenous expression of the mRNA of the three hypoxia-inducible
 RT factor a-subunits HIF-1a, HIF2a and HIF3a in rat liver.";
 RL Biochem. J. 354:531-537 (2001).
 CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
 CC EMBL: AJ277827; CAB96611.1; -
 DR HSBP; Q99814; 1P97.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0004871; F:signal transducer activity; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR GO; GO:0007165; P:signal transduction; IEA.
 DR InterPro; IPR001092; HLH_basic.
 DR InterPro; IPR001067; NUC_translocat.
 DR InterPro; IPR001610; PAC.
 DR InterPro; IPR000014; PAS.
 DR Pfam; PF00010; HLH; 1.
 DR Pfam; PF00989; PAS; 1.
 DR PRINTS; PR00785; NCTRNLOCATR.
 DR SMART; SM00353; HLH; 1.
 DR SMART; SM00086; PAC; 1.
 DR SMART; SM00091; PAS; 2.
 DR PROSITE; PSS0112; PAS; 2.
 SQ SEQUENCE 662 AA; 72887 MW; AC9672E3340544010 CRC64;

Query Match 50.2%; Score 820.5; DB 2; Length 662;
 Best Local Similarity 68.1%; Pred. No. 1.6e-59;
 Matches 177; Conservative 9; Mismatches 31; Indels 43; Gaps 5;

QY 9 RSNTELRKESGRDARSRSQETEVLYQLAHTLPARGVSAHLDKASITRLTISYLRMR 68

Db 7 RSSTELREKSRDARSRSQETEVLYQLAHTLPFARVSAHLDKASIMRLTISLYRMR 66
 Qy 69 LCAAGKRGKRGATGRLPEGCGFRHGTBRGRHGLPVGKCQ-----QAPGPOSV 117
 Db 67 LCAAG-----ENQVKGAGEPILDACYLKALGFWVWLTBGMAY 106
 Qy 118 DLCSSTLIHNPTPGTNS-LELIGHSIFDPIHPCDQELDALTTPRNLSKKLEAPTER 176
 Db 107 -----LSENVSKHLGSQLLELIGHSIFDPIHPCDQELDALTTPRELSSKSEAAATGR 160
 Qy 177 HFSLRMSTLTSRGRTNLKAAATWKVLCSGHMAAYKPPAOTSPPAGSPRSEPPLOCLVLI 236
 Db 161 HFSLRMSTLTSRGRTNLKAAATWKVLCSGHMAAYKPPAOTSPPAGSPRSEPPLOCLVLI 220
 Qy 237 CEALP-----OLPFDGATL 251
 Db 221 CEALPHPASLEPPLGRGAF 240

RESULT 8

09Y2N7 PRELIMINARY; PRT; 667 AA.
 AC 09Y2N7;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
 DE Putative homolog of hypoxia inducible factor three alpha (Hypoxia-inducible factor-3 alpha).
 GN Name=HIF-3A;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=21458277; PubMed=11573933; DOI=10.1006/dbrc.2001.5659;
 RA Hara S., Hamada J., Kobayashi C., Kondo Y., Imura N.,
 RT "Expression and characterization of hypoxia-inducible factor (HIF)-3alpha in human kidney: suppression of HIF-mediated gene expression by HIF-3alpha."
 RL Biochem. Biophys. Res. Commun. 287:808-813(2001).
 CC -1. SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
 DR EMBL; AC007133; AAD22668.1; -
 DR EMBL; AB054067; BAB69689.1; -
 DR PIR; JC7771, JC7771.
 DR HSSP; Q16665; 1U0B.
 DR GeneW; HGNC:15825; HIF3A.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0004871; F:signal transducer activity; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR GO; GO:0007165; P:signal transduction; IEA.
 DR InterPro; IPR001092; HLH_Dact1.
 DR InterPro; IPR001610; PAC.
 DR InterPro; IPR000014; PAS.
 DR Pfam; PF00010; HLH; 1.
 DR Pfam; PF00989; PAS; 1.
 DR PRINTS; PR00785; NCTRNLOCATR.
 DR SMART; SM00353; HLH; 1.

DR SMART; SM00086; PAC; 1.
 DR SMART; SM00091; PAS; 2.
 DR PROSITE; PS00112; PAS; 2.
 SQ SEQUENCE 667 AA; 72404 MW; 67B8794FF9DCCF4B CRC64;

Query Match 49.0%; Score 802; DB 2; Length 667;
 Best Local Similarity 68.0%; Pred. No. 5.5e-58;
 Matches 172; Conservative 11; Mismatches 28; Indels 42; Gaps 5;

Qy 9 RSSTELREKSRDARSRSQETEVLYQLAHTLPFARVSAHLDKASIMRLTISLYRMR 68
 Db 7 RSSTELREKSRDARSRSQETEVLYQLAHTLPFARVSAHLDKASIMRLTISLYRMR 66
 Qy 69 LCAAGKRGKRGATGRLPEGCGFRHGTBRGRHGLPVGKCQ-----QAPGPOSV 117
 Db 67 LCAAG-----ENQVKGAGEPILDACYLKALGFWVWLTBGMAY 106
 Qy 118 DLCSSTLIHNPTPGTNS-LELIGHSIFDPIHPCDQELDALTTPRNLSKKLEAPTER 176
 Db 107 -----LSENVSKHLGSQLLELIGHSIFDPIHPCDQELDALTTPQOTLSRKVEAPTER 160
 Qy 177 HFSLRMSTLTSRGRTNLKAAATWKVLCSGHMAAYKPPAOTSPPAGSPRSEPPLOCLVLI 236
 Db 161 CFSLRMSTLTSRGRTNLKAAATWKVLCSGHMAAYKPPAOTSPPAGSPRSEPPLOCLVLI 220
 Qy 237 CEALPQLPFDGA 249
 Db 221 CEALP-----HPGS 229

RESULT 9

09HA12 PRELIMINARY; PRT; 648 AA.
 ID 09HA12;
 AC 09HA12;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Hypothetical protein FJ11591.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Whole embryo;
 RX PubMed=14702039; DOI=10.1038/ng1285;
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R., Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H., Sekine M., Oobayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S., Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M., Nimomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoa S., Chiba Y., Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Horita T., Kusano J., Kanehori K., Takahashi Fujii A., Hara H., Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N., Mitsuhashi K., Yuki H., Oshino H., Sasaki N., Aotsuka S., Yoshihawa K., Matsunawa H., Ichihara T., Shiohara T., Sano S., Moriya S., Momiya A., Satoh N., Takami S., Terashima F., Wajake H., Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wajake H., Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B., Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y., Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujisawa T., Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S., Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsunura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M., Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T., Nakagawa K., Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K., Okumura K., Nagase T., Nomura N., Kikuchi H., Maehuo Y., Yamashita R.,

RA Nakai K., Yada T., Nakamura Y., Ohara O., Isozaki T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
cDNAs.";
RL Nat. Genet. 36:40-45(2004).
RA EMBL: AK021653; BAB13865.1; -.
DR HSSP: Q16665; ILOB.
DR GO: GO:0004871; F: signal transducer activity; IEA.
DR GO: GO:0006355; P: regulation of transcription, DNA-dependent; IEA.
DR GO: GO:0007165; P: signal transduction; IEA.
DR InterPro: IPR001610; PAC.
DR InterPro: IPR000014; PAS.
DR Pfam: PF00989; PAS; 1.
DR SMART: SMO0086; PAC; 1.
DR SMART: SMO0091; PAS; 2.
DR PROSITE: PS50112; PAS; 2.
SQ SEQUENCE 648 AA; 69995 MW; BBEFC744BC3F148B CRC64;

Query March 35.0%; Score 573; DB 2; Length 648;
Best Local Similarity 61.0%; Pred. No. 5.4e-39;
Matches 125; Conservative 11; Mismatches 27; Indels 42; Gaps 5;

QY 57 MRLTISYLRMRRLCAAGKGRATGRLPEGGGFRHGRHGRHGVKQC----- 109
DB 1 MRLTISYLRMRRLCAAG-----EMNGVAGSGEPDLACVYKALEGF 40

QY 110 ----QAPGPGVDLCSSSLIHNPTPGTNFSLIIGHISIPDFIHPCCQBEIQDALTFRPN 164
DB 41 VMVLTAGDMVY-----LSENVSKHGLSQLELIGHISIPDFIHPCCQBEIQDALTFRPQT 94

QY 165 LSKKKLEAPTEHRHPSLRMKSFTLSRGRTLNKAAATWVYHCSGHRAYKPPAOCSPAGSP 224
DB 95 LSRKKVLEAPTEHRCSLRMKSFTLSRGRTLNKAAATWVYHCSGHRAYKPPAOCSPAGSP 154

QY 225 RSEPPLOCLVLCERAIPLQLPFHGA 249
DB 155 DSEPPLOCLVLCERAIPLQLPFHGA 175

RESULT 10
ID Q9HAM5 PRELIMINARY; PRT; 199 AA.
AC Q9HAM5;
DT 01-MAR-2001 (Tremblrel. 16, Last Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Hypothetical protein Faj1359.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Whole embryo;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Ohtsushi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
RA Niinomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hiraoa S., Chiba Y., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
RA Togashi S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA Musasino K., Yuuki H., Oshima A., Sasaki N., Aocuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,

RA Fujimori Y., Komiyama M., Taehiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Sena T.,
RA Matsunura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togaishi T., Oyama M., Hara H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Maehuo Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isozaki T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
cDNAs.";
RL Nat. Genet. 36:40-45(2004).
RA EMBL: AK021421; BAB13819.1; -.
DR HSSP: Q99814; IP97.
SQ SEQUENCE 199 AA; 21598 MW; 646CB066A09B47C3 CRC64;

Query March 31.7%; Score 518.5; DB 2; Length 199;
Best Local Similarity 68.0%; Pred. No. 4.5e-35;
Matches 104; Conservative 9; Mismatches 13; Indels 27; Gaps 2;

QY 120 CSSSLIHNPTPGTNF-----STELIGHISIPDFIHPCCQBEIQ 156
DB 14 CTSWLTGCPSPASAPPTWRLSCASPSATCACTASAPQLEIGHISIPDFIHPCCQBEIQ 73

QY 157 DALTPRNLKSLKLEAPTEHRHPSLRMKSFTLSRGRTLNKAAATWVYHCSGHRAYKPPA 216
DB 74 DALTPQTLSSRRVLEAPTEHRCSLRMKSFTLSRGRTLNKAAATWVYHCSGHRAYKPPA 133

QY 217 QTSPPAGSPSEPPLOCLVLCERAIPLQLPFHGA 249
DB 134 QTSPPAGSPSEPPLOCLVLCERAIPLQLPFHGA 162

RESULT 11
ID H1FA CHICK STANDARD; PRT; 811 AA.
AC Q9YIB9;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Hypoxia-inducible factor 1 alpha (HIF-1 alpha) (HIF1 alpha).
GN Name=H1FA;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phaethonidae; Phaethonidae;
OC Gallus.
NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RA Takahashi T.;
RT "Molecular cloning and expression of an avian cDNA for hypoxia-
inducible factor-1 alpha in embryonic ventricular myocytes.";
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Functions as a master transcriptional regulator of the
adaptive response to hypoxia. Binds to core DNA sequence 5'-
[AG]CGTG-3' within the hypoxia response element (HRE) of target
gene promoters. Activation requires recruitment of transcriptional
coactivators (By similarity).
CC -1- SUBUNIT: Efficient DNA binding requires heterodimerization of an
alpha and a beta/ARNT subunit (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic in normoxia, nuclear
translocation in response to hypoxia (By similarity).
CC -1- DOMAIN: Contains two independent C-terminal transactivation
domains, NTRD and CTAD, which function synergistically. Their
transcriptional activity is repressed by an intervening inhibitory
domain (ID) (By similarity).
CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC -1- SIMILARITY: Contains 2 PAS (PER-ARNT-SIM) dimerization domains.
CC -1- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; AB013746; BAA34234.2; -.
 DR HSSP; Q16665; 1H2K.
 DR InterPro; IPR001092; HLH basic.
 DR InterPro; IPR001321; Hypoxindf1A.
 DR InterPro; IPR001610; PAC.
 DR InterPro; IPR000014; PAS.
 DR Pfam; PF00010; HLH; 1.
 DR Pfam; PF00785; PAC; 1.
 DR Pfam; PF00989; PAS; 2.
 DR PRINTS; PR01080; HYPOXIA1FLA.
 DR SMART; SM00353; HLH; 1.
 DR SMART; SM00086; PAC; 1.
 DR SMART; SM00091; PAS; 2.
 DR PROSITE; PS50888; HLH; 1.
 DR PROSITE; PS50112; PAS; 2.
 KW Activator; DNA-binding; Nuclear protein; Repeat;
 KW Transcription regulation.
 FT DNA_BIND 17 30 Basic motif.
 FT DOMAIN 31 71 Helix-loop-helix motif.
 FT DOMAIN 80 157 PAS 1.
 FT DOMAIN 228 298 PAS 2.
 FT DOMAIN 302 345 PAC.
 FT DOMAIN 401 587 ODD.
 FT DOMAIN 529 573 NTAD.
 FT DOMAIN 576 785 ID.
 FT DOMAIN 703 706 Nuclear localization signal (Potential).
 FT DOMAIN 718 721 Nuclear localization signal (Potential).
 FT DOMAIN 771 811 CTAD.
 FT DOMAIN 583 588 Poly-Ser.
 SQ SEQUENCE 811 AA; 90542 MW; D14CD9FC98F064CB CRC64;
 Query Match 30.7%; Score 503; DB 1; Length 811;
 Best Local Similarity 49.8%; Pred. No. 4.6e-33;
 Matches 119; Conservative 20; Mismatches 72; Indels 28; Gaps 5;
 QY 9 RSNTELREKSRDARRSRQSEYVLYOLATLPARQVSAHLDAKASIMRLTISYLRMR 68
 DB 12 RISSERRERKSRDARRSRQSEYVLYOLATLPARQVSAHLDAKASIMRLTISYLRMR 71
 QY 69 LCAAGGKRGKGRAT-----GRLLPEGGGGRHGRHGRGLGVCKQQAQPGQSYDLCS 123
 DB 72 LLDGSELTEANMEKELNCFYKALDGFVWVLSBDG-----DMYWS 113
 QY 124 LIHNPTPG-TNPSLELIGHSIFDFHPCDOBELDOLTPRPNLSSKKLEAFTERHPSLRM 182
 DB 114 ENWKKMGKLTQF--DLTGHSVFDFTPHCDHELRMLTHRNGPVKKGKQNTKTSFLIRM 171
 QY 183 KSLTTSRGRTNLKAATWVKVLHCSGHRMAVYPPAQTSPAGSPREBPQLCVLICEALP 241
 DB 172 KCLTTSRGRTNLKAATWVKVLHCSGHRMAVYPPAQTSPAGSPREBPQLCVLICEALP 228
 RESULT 12
 HIFA BOVIN STANDARD; PRT; 823 AA.
 ID HIFA BOVIN
 AC O9XTAS;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 05-JUN-2004 (Rel. 44, Last annotation update)
 DE Hypoxia-inducible factor 1 alpha (HIF-1 alpha) (HIF1 alpha).
 GN Name=HIF1A;
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovine; Bos
 NC NCB1_TaxID=9913;
 RN [1]

RP SEQUENCE FROM N.A.
 RC TISSUE=Artery;
 RX MEDLINE=99255430; PubMed=10320777; DOI=10.1016/S0167-4781(99)00048-2;
 RA Hara S., Kobayashi C., Imura N.;
 RT "Molecular cloning of cDNAs encoding hypoxia-inducible factor (HIF)-
 RL 1alpha and -2alpha of bovine arterial endothelial cells.";
 RL Biochim. Biophys. Acta 1445:237-243(1999).
 CC -1- FUNCTION: Functions as a master transcriptional regulator of the
 CC adaptive response to hypoxia. Under hypoxic conditions activates
 CC the transcription of over 40 genes, including, erythropoietin,
 CC glucose transporters, glycolytic enzymes, vascular endothelial
 CC growth factor, and other genes whose protein products increase
 CC oxygen delivery or facilitate metabolic adaptation to hypoxia.
 CC Plays an essential role in embryonic vascularization, tumor
 CC angiogenesis and pathophysiology of ischemic disease. Binds to
 CC core DNA sequence 5'-(AG)CGTG-3' within the hypoxia response
 CC element (HRE) of target gene promoters. Activation requires
 CC recruitment of transcriptional coactivators such as CREBBP and
 CC EP300. Activity is enhanced by interaction with both, NCOA1 or
 CC NCOA2. Interaction with redox regulatory protein APEX seems to
 CC activate CTAD and potentiates activation by NCOA1 and CREBBP (By
 CC similarity).
 CC -1- SUBUNIT: Efficient DNA binding requires heterodimerization of an
 CC alpha and a beta/ARNT subunit. Binds to the TAZ-type 1 domain of
 CC CREBBP and EP300. Interacts with NCOA1, NCOA2, APEX and HSP90 (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic in normoxia, nuclear
 CC translocation in response to hypoxia (By similarity).
 CC -1- DOMAIN: Contains two independent C-terminal transactivation
 CC domains, NTAD and CTAD, which function synergistically. Their
 CC transcriptional activity is repressed by an intervening inhibitory
 CC domain (ID) (By similarity).
 CC -1- PTM: In normoxia, is hydroxylated on Pro-402 and Pro-564 in the
 CC EGN1/PHD2. EGN1/PHD2 has also been shown to hydroxylate Pro-564.
 CC The hydroxylated prolines promote interaction with VHL, initiating
 CC rapid ubiquitination and subsequent proteasomal degradation. Under
 CC hypoxia, proline hydroxylation is impaired and ubiquitination is
 CC attenuated, resulting in stabilization (By similarity).
 CC -1- PTM: In normoxia, is hydroxylated on Asn-800 by HIF1AN, thus
 CC abrogating interaction with CREBBP and EP300 and preventing
 CC transcriptional activation (By similarity).
 CC -1- PTM: S-nitrosylated. All 15 free thiol groups are subjected to S-
 CC nitrosylation in vitro, however not all thiol groups seem to be
 CC nitrated in vivo (By similarity).
 CC -1- PTM: Acetylation of Lys-532 by ARD1 increases interaction with VHL
 CC and stimulates subsequent proteasomal degradation (By similarity).
 CC -1- PTM: Requires phosphorylation for DNA-binding (By similarity).
 CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
 CC -1- SIMILARITY: Contains 2 PAS (PER-ARNT-SIM) dimerization domains.
 CC -1- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; AB018398; BAA78675.1; -.
 DR HSSP; Q16665; 1H2K.
 DR InterPro; IPR001092; HLH basic.
 DR InterPro; IPR001321; Hypoxindf1A.
 DR InterPro; IPR001610; PAC.
 DR InterPro; IPR000014; PAS.
 DR Pfam; PF00010; HLH; 1.
 DR Pfam; PF00785; PAC; 1.
 DR Pfam; PF00989; PAS; 2.
 DR PRINTS; PR01080; HYPOXIA1FLA.
 DR SMART; SM00353; HLH; 1.
 DR SMART; SM00086; PAC; 1.
 DR SMART; SM00091; PAS; 2.

DR PROSITE; PSS0888; HLH; 1.
 DR PROSITE; PSS0112; PAS; 2.
 KW Acetylation; Activator; DNA-binding; Hydroxylation; Nuclear protein;
 KW Phosphorylation; Repeat; S-nitrosylation; Transcription regulation.
 FT DOMAIN 718 721
 FT DNA_BIND 17 30
 FT DOMAIN 31 71
 FT DOMAIN 85 158
 FT DOMAIN 228 298
 FT DOMAIN 302 345
 FT DOMAIN 401 600
 FT DOMAIN 531 575
 FT DOMAIN 576 782
 FT DOMAIN 783 823
 FT DOMAIN 715 718
 FT MOD_RES 90 90
 FT MOD_RES 139 139
 FT MOD_RES 173 173
 FT MOD_RES 194 194
 FT MOD_RES 210 210
 FT MOD_RES 219 219
 FT MOD_RES 224 224
 FT MOD_RES 255 255
 FT MOD_RES 334 334
 FT MOD_RES 337 337
 FT MOD_RES 359 359
 FT MOD_RES 402 402
 FT MOD_RES 520 520
 FT MOD_RES 532 532
 FT MOD_RES 564 564
 FT MOD_RES 755 755
 FT MOD_RES 777 777
 FT MOD_RES 797 797
 FT MOD_RES 800 800
 SQ SEQUENCE 823 AA; 92127 MW; 12674E467A61B1A1 CRC64;

Query Match 30.5%; Score 499; DB 1; Length 823;
 Best Local Similarity 46.6%; Pred. No. 1e-32;
 Matches 115; Conservative 26; Mismatches 58; Indels 48; Gaps 6;

QY 11 NTELRKESKSDAARSRSOETEVLYOLAHTLPARGVSAHLDKASIMRLTISYLRMRLC 70
 DB 14 SSKRKESKSDAARSRSKSESEVFEYLAHQLPHPNVSSHLDKASVRLTISYLRYKRL 73
 QY 71 AAGKGRATGRLLPEPGGFRGTHRGHGLPVGKCOA-----PGQSVYDLC 120
 DB 74 DAG-----DLIDEMKAQNMCFYLKALDGFVWLTD 105
 QY 121 SSSLIH-----NPTPG-TNPSLELIGHSIPDFIHPCDQELQDALTPRPNLSSKKLEAP 174
 DB 106 DGDMIYISDNVNMKMGITQF--ELTGHSVDFTHPCDHEMRMLTHRNGLVKKGEQNT 163
 QY 175 ERHPSLMMKSTLTSGRTINLKAATWKLHCSGMRAYKPPAOTSPAGSPRSEPPLOCLV 234
 DB 164 QRSFPLMKCTLTSGRTINLKAATWKLHCSGMRAYKPPAOTSPAGSPRSEPPLOCLV 221
 QY 235 LICEAIP 241
 DB 222 LICEPIP 228

RESULT 13
 Q6IV47 PRELIMINARY; PRT; 823 AA.
 ID O6IV47;
 AC O6IV47;
 DT 05-JUL-2004 (TREMBLrel. 27, Created)
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
 DE Hypoxia-inducible factor-1a.
 GN Name=HIF-1A;
 OS Bos mutus graminis (Yak).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Mammalia;

OC Bovinae; Bos.
 OX NCBI_TaxID=30521;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Dolt K.S., Qadar Pasha M.A.;
 RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
 DR EMBL; AY621118; AAT39520.1; -;
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0004871; F:signal transducer activity; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR GO; GO:0007165; P:signal transduction; IEA.
 DR InterPro; IPR001092; HLH_Basic.
 DR InterPro; IPR001610; PAC.
 DR InterPro; IPR000014; PAS.
 DR Pfam; PF00010; HLH; 1.
 DR Pfam; PF00785; PAC; 1.
 DR Pfam; PF00989; PAS; 2.
 DR PRINTS; PR01080; HYPOXIA1F1A.
 DR SMART; SM00353; HLH; 1.
 DR SMART; SM00086; PAC; 1.
 DR SMART; SM00091; PAS; 2.
 DR PROSITE; PSS0888; HLH; 1.
 DR PROSITE; PSS0112; PAS; 2.
 SQ SEQUENCE 823 AA; 92128 MW; A6E388E4FEA15705 CRC64;

Query Match 30.5%; Score 499; DB 2; Length 823;
 Best Local Similarity 46.6%; Pred. No. 1e-32;
 Matches 115; Conservative 26; Mismatches 58; Indels 48; Gaps 6;

QY 11 NTELRKESKSDAARSRSOETEVLYOLAHTLPARGVSAHLDKASIMRLTISYLRMRLC 70
 DB 14 SSKRKESKSDAARSRSKSESEVFEYLAHQLPHPNVSSHLDKASVRLTISYLRYKRL 73
 QY 71 AAGKGRATGRLLPEPGGFRGTHRGHGLPVGKCOA-----PGQSVYDLC 120
 DB 74 DAG-----DLIDEMKAQNMCFYLKALDGFVWLTD 105
 QY 121 SSSLIH-----NPTPG-TNPSLELIGHSIPDFIHPCDQELQDALTPRPNLSSKKLEAP 174
 DB 106 DGDMIYISDNVNMKMGITQF--ELTGHSVDFTHPCDHEMRMLTHRNGLVKKGEQNT 163
 QY 175 ERHPSLMMKSTLTSGRTINLKAATWKLHCSGMRAYKPPAOTSPAGSPRSEPPLOCLV 234
 DB 164 QRSFPLMKCTLTSGRTINLKAATWKLHCSGMRAYKPPAOTSPAGSPRSEPPLOCLV 221
 QY 235 LICEAIP 241
 DB 222 LICEPIP 228

RESULT 14
 HIFA_RAT STANDARD; PRT; 825 AA.
 ID HIFA_RAT
 AC O35800; Q9WTU9;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Hypoxia-inducible factor 1 alpha (HIF-1 alpha) (HIF1 alpha).
 GN Name=Hif1a;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wistar; TISSUE=Hepatocytes;
 RC MEDLINE=2114367; PubMed=11237857; DOI=10.1042/0264-6021:3540531;
 RA Kietzmann T., Cornesse Y., Brechtel K., Modaresi S., Jungermann K.;
 RT "periteneous expression of the mRNA of the three hypoxia-inducible factor a-subunits HIF-1a, HIF2a and HIF3a in rat liver.";


```
RESULT 15
Q6EMI3 PRELIMINARY; PRT; 489 AA.
ID Q6EMI3
AC Q6EMI3;
DT 25-OCT-2004 (TRENBLREL. 28, Created)
DT 25-OCT-2004 (TRENBLREL. 28, Last sequence update)
DT 25-OCT-2004 (TRENBLREL. 28, Last annotation update)
DE Hypoxia-inducible factor 1 alpha (Fragment).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX PubMed=15162502; DOI=10.1002/dvdy.20049;
RA Sipe C.W., Gruber E.J., Saha M.S.;
RT "Short upstream region drives dynamic expression of hypoxia-inducible
factor 1alpha during Xenopus development.";
RL Dev. Dyn. 230:229-238(2004).
CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
DR EMBL; AY189821; AAC72733.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR01092; HLH basic.
DR InterPro; IPR01321; Hypoxindf1A.
DR InterPro; IPR01067; Nuc_translocat.
DR InterPro; IPR01610; PAC.
DR InterPro; IPR00014; PAS.
DR Pfam; PF00010; HLH; 1.
DR Pfam; PF00785; PAC; 1.
DR Pfam; PF00989; PAS; 2.
DR PRINTS; PR01080; HYPOXIA1F1A.
DR PRINTS; PR00785; NCTRNLOCATR.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00086; PAC; 1.
DR SMART; SM00091; PAS; 2.
DR PROSITE; PS50888; HLH; 1.
DR PROSITE; PS50112; PAS; 2.
FT NON_TER 1
FT NON_TER 489
SQ SEQUENCE 489 AA; 55613 MW; 39E1ACA3EA6AF766 CRC64;

Query Match 30.3%; Score 495; DB 2; Length 489;
Best Local Similarity 48.6%; Pred. No. 1.2e-32;
Matches 114; Conservative 27; Mismatches 71; Indels 18; Gaps 6;

QY 13 ELRREKSRDAARSRROETEVLYQLAHTLPPARGVSAHLDKASIMRLTISYLRMRLCAA 72
DB 1 ERREKSRDARCRSRKSESVFYELSHLPLPHNVSSHLDKASIMRLAISYLRRLDLA 60
QY 73 GGRKGRATGRLLPBGPGFRGTRGRGRHGLPVGKCOAPGPQSVYDLCSSSLINPTPG- 131
DB 61 GDLDGETD--LDQLNCF---YLKALEGFVLVLTBERG-----DWIYSENVNKKMGGL 107
QY 132 TNFSLLELIGHSIFPFIRPCDOEELODATPPRNLSKKKLGAFTERRHPSLRKSTLTSRGR 191
DB 108 TGF--ELTGHSVFPTFTRCDHEELREMLTFRNGPAKKGKEQITERSFFLRKKTCTVTSRGR 165
QY 192 TLNLKAATWKLVLHSGHMRAYKPPAQTSPAGSPRSEPPLOCLVLICEAIP 241
DB 166 TVNIKSAITWKLVLHCTGHMRVYDNNANQKHG--YKRPKCMVVICPIP 213
```

Search completed: June 15, 2005, 16:00:09
Job time : 176 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 16:00:19 ; Search time 157 Seconds
(without alignments)
749.578 Million cell updates/sec

Title: US-09-896-791B-3

Perfect score: 1636
Sequence: 1 MALGGLQVRNTELRKESR.....TESSLPSWVLMALNRKNCPG 307

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1710399 seqs, 38333425 residues

Total number of hits satisfying chosen parameters: 1710399

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA: *
1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubppaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubppaa/US10E_PUBCOMB.pep:*
18: /cgn2_6/ptodata/1/pubppaa/US10F_PUBCOMB.pep:*
19: /cgn2_6/ptodata/1/pubppaa/US10G_PUBCOMB.pep:*
20: /cgn2_6/ptodata/1/pubppaa/US10H_PUBCOMB.pep:*
21: /cgn2_6/ptodata/1/pubppaa/US10I_PUBCOMB.pep:*
22: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1636	100.0	307	9	US-09-896-791B-3
2	1278	78.1	324	15	US-10-307-828A-20
3	795	48.6	705	14	US-10-154-386-2
4	498	30.4	827	10	US-09-919-039-149
5	498	30.4	827	14	US-10-247-671-137
6	495	30.3	623	10	US-09-967-388-2
7	495	30.3	735	16	US-10-032-361-2
8	495	30.3	826	9	US-09-922-958-4
9	495	30.3	826	9	US-09-833-790-235
10	495	30.3	826	9	US-09-736-457-330
11	495	30.3	826	9	US-09-902-941-330

12	495	30.3	826	9	US-09-849-626-330	Sequence 330, App
13	495	30.3	826	10	US-09-967-388-4	Sequence 4, Appl1
14	495	30.3	826	10	US-09-476-300-330	Sequence 330, App
15	495	30.3	826	13	US-10-028-158-23	Sequence 23, Appl
16	495	30.3	826	13	US-10-101-812-13	Sequence 10, Appl
17	495	30.3	826	14	US-10-101-662A-9	Sequence 9, Appl1
18	495	30.3	826	14	US-10-101-816-2	Sequence 2, Appl1
19	495	30.3	826	14	US-10-101-816-5	Sequence 5, Appl1
20	495	30.3	826	14	US-10-101-816-6	Sequence 6, Appl1
21	495	30.3	826	14	US-10-101-816-7	Sequence 7, Appl1
22	495	30.3	826	14	US-10-017-954-330	Sequence 330, App
23	495	30.3	826	14	US-10-115-987B-14	Sequence 14, Appl
24	495	30.3	826	14	US-10-287-670-9	Sequence 9, Appl1
25	495	30.3	826	14	US-10-113-872-330	Sequence 330, App
26	495	30.3	826	14	US-10-423-419-2	Sequence 2, Appl1
27	495	30.3	826	15	US-10-283-017-330	Sequence 330, App
28	495	30.3	826	16	US-10-032-361-1	Sequence 1, Appl1
29	495	30.3	826	16	US-10-854-483-2	Sequence 2, Appl1
30	495	30.3	826	16	US-10-854-483-5	Sequence 5, Appl1
31	495	30.3	826	16	US-10-854-483-6	Sequence 6, Appl1
32	495	30.3	826	16	US-10-854-483-7	Sequence 7, Appl1
33	495	30.3	826	16	US-10-859-935-10	Sequence 10, Appl
34	495	30.3	826	16	US-10-859-935-23	Sequence 23, Appl
35	495	30.3	826	17	US-10-901-583-18	Sequence 18, Appl
36	491	30.0	823	14	US-10-205-342-13	Sequence 13, Appl
37	467	28.5	870	14	US-10-101-816-4	Sequence 4, Appl1
38	467	28.5	870	16	US-10-854-483-4	Sequence 4, Appl1
39	466	28.5	485	14	US-10-121-235-4	Sequence 4, Appl1
40	466	28.5	870	14	US-10-121-235-6	Sequence 6, Appl1
41	281	17.2	1507	14	US-10-080-608A-37	Sequence 37, Appl
42	281	17.2	1507	15	US-10-370-685-126	Sequence 126, App
43	278.5	17.0	901	16	US-10-408-765A-1259	Sequence 1259, App
44	276.5	16.9	903	16	US-10-408-765A-1332	Sequence 1332, App
45	256	15.6	570	9	US-09-923-684-4	Sequence 4, Appl1

ALIGNMENTS

RESULT 1
US-09-896-791B-3
Sequence 3, Application US/09896791B
Patent NO. US20020165140A1
GENERAL INFORMATION:
APPLICANT: Berkenstam, Anders
APPLICANT: Poellinger, Lorenz
TITLE OF INVENTION: SCREENING METHODS
FILE REFERENCE: 13425-040001
CURRENT APPLICATION NUMBER: US/09/896, 791B
CURRENT FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: 60/217,570
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: SE 0002551-0
PRIOR FILING DATE: 2000-07-06
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 307
TYPE: PRT
ORGANISM: Mus musculus
US-09-896-791B-3

Query Match 100.0%; Score 1636; DB 9; Length 307;
Best Local Similarity 100.0%; Pred. No. 6.3e-146; Indels 0; Gaps 0;
Matches 307; Conservative 0; Mismatches 0;
QY 1 MALGGLQVRNTELRKESRDAARRSRQSTEVLYQLAHTLPFARGVSAHLDKASINRLT 60
DB 1 MALGGLQVRNTELRKESRDAARRSRQSTEVLYQLAHTLPFARGVSAHLDKASINRLT 60
QY 61 ISYRMRLCAAGKGRATGRLLPBGPGFRRGTHRRGRHGLPVGCKQAPGQSYDLC 120

```

Db      61 ISYRMHRLCAAGKGRGATGRLLPEGPGGFRHGTNRGRHGLPVGKCOQAPGQSVLDC 120
Qy      121 SSSLINHPPTGTFNLSLEIGHISFDPIHPCDQELQDALTPRPNLSSKKLEAPTERHPSL 180
Db      121 SSSLINHPPTGTFNLSLEIGHISFDPIHPCDQELQDALTPRPNLSSKKLEAPTERHPSL 180
Qy      181 RMKSTLTSRGRITNLKAATWKVHLHCSGHRAYKPPAOTSPPAGSPRSEPPLOCLVLI CEAI 240
Db      181 RMKSTLTSRGRITNLKAATWKVHLHCSGHRAYKPPAOTSPPAGSPRSEPPLOCLVLI CEAI 240
Qy      241 POLPFHGGATLGLPQEKTPISTLTFTPLMKALLCLVKRPVQVLOGKGTSSLPSPVLMAL 300
Db      241 POLPFHGGATLGLPQEKTPISTLTFTPLMKALLCLVKRPVQVLOGKGTSSLPSPVLMAL 300
Qy      301 NRKNCPG 307
Db      301 NRKNCPG 307

RESULT 2
US-10-307-928A-20
; Sequence 20, Application US/10307928A
; Publication No. US20030229016A1
; GENERAL INFORMATION:
; APPLICANT: Alsebrook, John P.
; APPLICANT: Anderson, David W.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Catterton, Elna
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Gorman, Linda
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Ji, Weizhen
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Li, Li
; APPLICANT: Paturajan, Meera
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Vermet, Corine A.M.
; APPLICANT: Voess, Edward Z.
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYPEPTIDES ENCODING THEM AND METHODS C
; TITLE OF INVENTION: THE SAME
; FILE REFERENCE: 24102-502D
; CURRENT APPLICATION NUMBER: US/10/307,928A
; PRIOR FILING DATE: 2002-12-02
; PRIOR APPLICATION NUMBER: 60/341,477
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/341,540
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/342,592
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/344,903
; PRIOR FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/373,288
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/380,981
; PRIOR FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: 60/381,495
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/383,744
; PRIOR FILING DATE: 2002-05-28
; PRIOR APPLICATION NUMBER: 60/384,024
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: 60/401,788
; PRIOR FILING DATE: 2002-08-07
; Remaining Prior Application data removed - See File Wrapper or PAM.
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: Curaseqlist version 0.1
; SEQ ID NO 20
; LENGTH: 324
; TYPE: PRT

```

```

; ORGANISM: Homo sapiens
US-10-307-928A-20
Query Match 78.1%; Score 1278; DB 15; Length 324;
Best Local Similarity 79.9%; Pred. No. 4,6e-112;
Matches 250; Conservative 14; Mismatches 41; Indels 8; Gaps 3;

Qy      1 MALGLQVRNTELRKESRDARSRRSQTEVLYOLAHTLPARGVSAHLDAKASIMRLT 60
Db      1 MALGLQVRNTELRKESRDARSRRSQTEVLYOLAHTLPARGVSAHLDAKASIMRLT 60
Qy      61 ISYRMHRLCAAGKGRGATGRLLPEGPGGFRHGTNRGRHGLPVGKCOQAPGQSVLDC 120
Db      61 ISYRMHRLCAAGKGRGATGRLLPEGPGGFRHGTNRGRHGLPVGKCOQAPGQSVLDC 120
Qy      121 SSSLINHPPTGTFNLSLEIGHISFDPIHPCDQELQDALTPRPNLSSKKLEAPTERHPSL 180
Db      121 SSSLINHPPTGTFNLSLEIGHISFDPIHPCDQELQDALTPRPNLSSKKLEAPTERHPSL 180
Qy      181 RMKSTLTSRGRITNLKAATWKVHLHCSGHRAYKPPAOTSPPAGSPRSEPPLOCLVLI CEAI 240
Db      181 RMKSTLTSRGRITNLKAATWKVHLHCSGHRAYKPPAOTSPPAGSPRSEPPLOCLVLI CEAI 240
Qy      241 POLPFHGGATLGLPQEKTPISTLTFTPLMKALLCLVKRPVQVLOGKGTSSLPSPVLMAL 300
Db      241 POLPFHGGATLGLPQEKTPISTLTFTPLMKALLCLVKRPVQVLOGKGTSSLPSPVLMAL 300
Qy      295 WVLALNRKNCPG 307
Db      295 EVWMALNGKNCAG 311

RESULT 3
US-10-154-386-2
; Sequence 2, Application US/10154386
; Publication No. US20030026793A1
; GENERAL INFORMATION:
; APPLICANT: Angiogene Inc.
; APPLICANT: Guy, Louis-Georges
; TITLE OF INVENTION: HIPPOXIA INDUCING FACTORS AND USES THEREOF FOR INDUCING ANGIOGENESIS
; TITLE OF INVENTION: IMPROVING MUSCULAR FUNCTIONS
; FILE REFERENCE: 5600-81
; CURRENT APPLICATION NUMBER: US/10/154,386
; PRIOR FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/292,630
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/354529
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 705
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-154-386-2
Query Match 48.6%; Score 795; DB 14; Length 705;
Best Local Similarity 67.6%; Pred. No. 5,6e-66;
Matches 171; Conservative 11; Mismatches 29; Indels 42; Gaps 5;

Qy      9 RSTTELRKESRDARSRRSQTEVLYOLAHTLPARGVSAHLDAKASIMRLTISYLMR 68
Db      3 RSTTELRKESRDARSRRSQTEVLYOLAHTLPARGVSAHLDAKASIMRLTISYLMR 62
Qy      69 LCAAGKGRGATGRLLPEGPGGFRHGTNRGRHGLPVGKCOQAPGQSV 117
Db      63 LCAAG-----EMQVAGAGEPLDACYLKALGFVWVLTABGDMAY 102
Qy      118 DLSSSLINHPPTGTFNLSLEIGHISFDPIHPCDQELQDALTPRPNLSSKKLEAPTER 176
Db      103 -----LSENVSGHGLGSELEIGHISFDPIHPCDQELQDALTPQOTLSRRVBAFTER 156
Qy      177 HPSLRMKSTLTSRGRITNLKAATWKVHLHCSGHRAYKPPAOTSPPAGSPRSEPPLOCLVLI 236

```

Db	157	CFSLRMKSTLISRGRTINLKAATWKLKYNCSGHRAYKPAQGISPASSPDSSEPLQCPVLI	216
Qy	237	CEAIPQLPEFHDA	249
		:	:
Db	217	CEAIP---HPGS	225

```

RESULT 4
US-09-919-039-149
; Sequence 149, Application US/09919039
; Publication No. US20030108871A1
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 149
; LENGTH: 827
; TYPE: PR1
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030108871A1 1250434CD1
; US-09-919-039-149

```

```

RESULT 5
US-10-247-671-137
: Sequence 137, Application US/10247671
: Publication NO. US20030194721A1
: GENERAL INFORMATION:
: APPLICANT: Mikita, Thomas
: APPLICANT: Shiffman, Dov
: APPLICANT: Porter, Gordon, J.
: APPLICANT: Kaser, Matthew R.
: TITLE OR INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
: FILE REFERENCE: PA-0050 US
: CURRENT APPLICATION NUMBER: US/10/247,671
: CURRENT FILING DATE: 2002-09-18
: PRIOR APPLICATION NUMBER: 60/323,784
: PRIOR FILING DATE: 2001-09-19
: NUMBER OF SEQ. ID NOS: 186
: SOFTWARE: PERL Program

```

```

: SEQ ID NO 137
:
: LENGTH: 827
:
: TYPE: PRT
:
: ORGANISM: Homo sapiens
:
: FEATURE:
:
: NAME/KEY: misc feature
:
: OTHER INFORMATION: Incyte ID No. US20030194721A1 1250434CDD
:
: US-10-247-671-137

```

```

: RESULT 6
: US-09-967-388-2
: Sequence 2, Application US/09967388
: Publication No. US20030103956A1
: GENERAL INFORMATION:
: APPLICANT: JEFFERY M. ARBEIT
: TITLE OF INVENTION: USE OF HIF-1ALPHA VARIANTS TO ACCELERATE
: FILE REFERENCE: UC077.001A
: CURRENT APPLICATION NUMBER: US/09/967,388
: CURRENT FILING DATE: 2001-09-28
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 2
: LENGTH: 623
: TYPE: PRT
: ORGANISM: HUMAN
: US-09-967-388-2

```



```
US-09-736-457-330
; Sequence 330, Application US/09736457
; Patent No. US20020168637A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Tongtong
; APPLICANT: Bangur, Chaicanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 330
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-736-457-330

Query Match      30.3%; Score 495; DB 9; Length 826;
Best Local Similarity 46.2%; Pred. No. 1.5e-37;
Matches 114; Conservative 26; Mismatches 59; Indels 48; Gaps 6;

Cy 11 NTELRKESRDAAARRSRQSETEVLYQLAHTLPFARGVSAHLDRKASIMELTISYLRMRLC 70
Db 14 SSERKESKSDAARRSRKSESEVFEYELAHQLPPLHNVSHLDRKASVRLTISYLRVRL 73
Cy 71 AAGKRGATGRLLPBPGGFRHGTNRGRHGLPVGKCOA-----PGQSYDLC 120
Db 74 DAG-----DLIDEDMKAKQNCFFYLKALDGFVWVLT 105
Cy 121 SSLSLIH-----NPTPG-TNFSLELIGHSIPDFTHPCQOELODALTPRPNLKKKLEAPT 174
Db 106 DGDMTIYSDNVNKKYMGTLQF--ELTGHSVDFPTHPCHEEREMLTNRNGLVKKGKQNT 163
Cy 175 ERHPSLNMKSTLTSRGRTLNKATWVKVLCSCGHRAYKPPAOTSPPAGSPRSEBPLOCLV 234
Db 164 QRSFPLMKCTLTISRGRTNMKSAWKVLCGTHIHVDYDMSNPQCG--YKKEPMTCLV 221
Cy 235 LICEAIP 241
Db 222 LICEPIP 228

RESULT 11
US-09-902-941-330
; Sequence 330, Application US/09902941
; Patent No. US20020172952A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaicanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C17
; CURRENT APPLICATION NUMBER: US/09/902,941
; CURRENT FILING DATE: 2001-07-10
```

```
; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 330
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-902-941-330

Query Match      30.3%; Score 495; DB 9; Length 826;
Best Local Similarity 46.2%; Pred. No. 1.5e-37;
Matches 114; Conservative 26; Mismatches 59; Indels 48; Gaps 6;

Cy 11 NTELRKESRDAAARRSRQSETEVLYQLAHTLPFARGVSAHLDRKASIMELTISYLRMRLC 70
Db 14 SSERKESKSDAARRSRKSESEVFEYELAHQLPPLHNVSHLDRKASVRLTISYLRVRL 73
Cy 71 AAGKRGATGRLLPBPGGFRHGTNRGRHGLPVGKCOA-----PGQSYDLC 120
Db 74 DAG-----DLIDEDMKAKQNCFFYLKALDGFVWVLT 105
Cy 121 SSLSLIH-----NPTPG-TNFSLELIGHSIPDFTHPCQOELODALTPRPNLKKKLEAPT 174
Db 106 DGDMTIYSDNVNKKYMGTLQF--ELTGHSVDFPTHPCHEEREMLTNRNGLVKKGKQNT 163
Cy 175 ERHPSLNMKSTLTSRGRTLNKATWVKVLCSCGHRAYKPPAOTSPPAGSPRSEBPLOCLV 234
Db 164 QRSFPLMKCTLTISRGRTNMKSAWKVLCGTHIHVDYDMSNPQCG--YKKEPMTCLV 221
Cy 235 LICEAIP 241
Db 222 LICEPIP 228

RESULT 12
US-09-849-626-330
; Sequence 330, Application US/09849626
; Publication No. US20020197699A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaicanya
; APPLICANT: Fanger, Gary
; APPLICANT: Wang, Aijun
; APPLICANT: Wang, Tongtong
; APPLICANT: McNeel, Anne
; APPLICANT: McNeel, Patricia
; APPLICANT: Clapper, Jonathan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C16
; CURRENT APPLICATION NUMBER: US/09/849,626
; CURRENT FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 1926
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 330
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-849-626-330

Query Match      30.3%; Score 495; DB 9; Length 826;
Best Local Similarity 46.2%; Pred. No. 1.5e-37;
Matches 114; Conservative 26; Mismatches 59; Indels 48; Gaps 6;

Cy 11 NTELRKESRDAAARRSRQSETEVLYQLAHTLPFARGVSAHLDRKASIMELTISYLRMRLC 70
Db 14 SSERKESKSDAARRSRKSESEVFEYELAHQLPPLHNVSHLDRKASVRLTISYLRVRL 73
Cy 71 AAGKRGATGRLLPBPGGFRHGTNRGRHGLPVGKCOA-----PGQSYDLC 120
Db 74 DAG-----DLIDEDMKAKQNCFFYLKALDGFVWVLT 105
Cy 121 SSLSLIH-----NPTPG-TNFSLELIGHSIPDFTHPCQOELODALTPRPNLKKKLEAPT 174
Db 106 DGDMTIYSDNVNKKYMGTLQF--ELTGHSVDFPTHPCHEEREMLTNRNGLVKKGKQNT 163
```

QY	175	ERFESLRMSTLTSRGTINTLKAAITWVLCGSHRAYKPPAOTSPPASPSSEPLQCLV	234
	:	: :	:
Dd	164	QRSEFLRMCTLTLSRGTINIKSAITWVLCHCGIHLIVDTINSNPQC--YKRPMTCLV	221
QY	235	LICEATP	241
Dd	222	LICEPIIP	228

```

RESULT 13
US-09-967-388-4
: Sequence 4, Application US/09967388
: Publication No. US20030103956A1
GENERAL INFORMATION:
APPLICANT: JEFFEREY M. ARBEIT
: TITLE OF INVENTION: USE OF HIF-1ALPHA VARIANTS TO ACCELERATE
: FILE REFERENCE: WOUND HEALING
: CURRENT APPLICATION NUMBER: UC077.001A
: CURRENT FILING DATE: 2001-09-28
NUMBER OF SEQ ID NOS: 4
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 4
LENGTH: 826
:
: TYPE: PRT
: ORGANISM: HUMAN
: US-09-967-388-4

```

	Query Match	30.3%	Score 495;	DB 10;	Length 826;
	Best Local Similarity	46.2%;	Pred. No. 1.5e-37;		
	Matches	114;	Conservative	26;	Mismatches 59; Indels 48; Gaps 6;
CY	11 NTELRKESRPAASRRSQEETVLYOLAHITPPFARGVSAMHLDKASIMPLTISTYLEMRHLIC	70			
Db	14 SSEERKEKSRPAASRRKESVEFEVLAHQPLPHNVSSHLDDKASVMMLTISTYLVRKLTL	73			
CY	71 AAGGKRGRATGRLLPEGGPGFRHGTHRGHRGLPVGKCOQA-----PGPSVDLCC	120			
Db	74 DAG-----DLDIIDDMKAQNMNCFYLKALDGFVMVLTD	105			
CY	121 SSLLIH-----NPTPG--TNFSLELIHSIPDPFHPCOEELIQDALTPRPNSKKLTAAFT	174			
Db	106 DGDMTIVSDNVKMYGLTFQ--ELTGHSVFPDFTHPCDHEENREMLITHNGLVKKGKEONT	163			
CY	175 ERHNSLRNKSTLTLSRGRTLNILKAAITWKVLCSCGMRAVYPAQSPPASPSPSEPPLCLV	234			
Db	164 QRSFFLRMKCTLTLSRGRTNIKSAITWVKLHTCGIHHVDTNSNPQCQ--YKKPMTCLV	221			
CY	235 LICEAIP 241				
Db	222 LICBPIP 228				

```

! ORGANISM: HUMAN
US-09-967-388-4

Query Match      30.3%; Score 495; DB 10; Length 826;
Best Local Similarity 46.2%; Pred. No. 1.5e-37;
Matches 114; Conservative 26; Mismatches 59; Indels 48; Gaps 6;

QY 11 NTELRLKESRPAARSRSQSETEVLYOLAHLPFARGVSAHLDKASIMELTISYLMHRLC 70
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 14 SSERKKEKSRPAASRSRKESEVPEFLAHQPLPINVSHIDPKASVMELTISYLRKLL 73

QY 71 AAGKRGKRGATGRLLEPGGFRHGTTHRRGRHGLPVKGCQQA-----PGQSYDLC 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 74 DAG-----DLIEDMKKQMMNCFYLKALDGFVMVLTLD 105

QY 121 SSSLIH-----NPTPG-TNPSLELLGHSIPDPIHCDOEELDALTTPRPNLSKKULEAPT 174
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 106 DGDMLIYSDNVNKKYMGLTQF--ELTGHSVFPDTPHCDHEEMKEMLTHNGLVKKGEONT 163

QY 175 ERHFSLMKSTLTSRGRTALNKAATWKLHSGHMRARAKPPAQTSPAGSPRSEPLQCLV 234
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 164 QRSFPLNRKCLTISRGRMTNKSATWKLHCTGHITHVDTNSNQDQC--YKKPMTGLV 221

QY 235 LICEAIP 241
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 222 LICEPIP 228

```

```

RESULT 14
US-09-476-300-330
; Sequence 330 Application US/09476300
; Publication No. US20030125245A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Bangur, Chaltanya S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C3
; CURRENT APPLICATION NUMBER: US/09/476,300
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 785
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 330
; LENGTH: 826
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-476-300-330

```

	Query Match	30.3%	Score 495;	DB 13;	Length 826;
	Best Local Similarity	46.2%;	Pred. No. 1.5e-37;		
	Matches	114;	Conservative	26;	Mismatches 59; Indels 48; Gaps 6;
Cy	11	NTELREKESRDAAASRSBSQETEVLYQIAHTLPARGVASAHLDKASIMRLTTSYLMARLC	70		
		:::::			
		:::::			
		:::::			
		:::::			
		:::::			
		:::::			
		:::::			
		:::::			

Oy	235	LICRAIP	241
Db	222	LICRIP	228

Search completed: June 15, 2005, 16:13:41
Job time : 159 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 15:50:48 ; Search time 42 Seconds
(without alignments)
545.649 Million cell updates/sec

Title: US-09-896-791B-3

Perfect score: 1636
Sequence: 1 MALGLOKRVSRNTBLRKEKSR.....TESSLPSWLMALNRKNCPG 307

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	498	30.4	827	4	US-09-919-039-149
2	495	30.3	245	4	US-09-438-833-3
3	495	30.3	330	4	US-09-438-833-4
4	495	30.3	623	4	US-09-967-388-2
5	495	30.3	652	4	US-09-438-833-5
6	495	30.3	813	4	US-09-438-833-12
7	495	30.3	826	1	US-08-785-241-6
8	495	30.3	826	2	US-08-460-473B-2
9	495	30.3	826	3	US-08-915-213-2
10	495	30.3	826	3	US-09-148-547-2
11	495	30.3	826	3	US-09-235-217-2
12	495	30.3	826	3	US-09-380-662-23
13	495	30.3	826	4	US-09-438-833-1
14	495	30.3	826	4	US-09-702-705-330
15	495	30.3	826	4	US-09-736-457-330
16	495	30.3	826	4	US-09-383-581-2
17	495	30.3	826	4	US-09-614-124B-330
18	495	30.3	826	4	US-09-671-325-330
19	495	30.3	826	4	US-09-589-184-330
20	495	30.3	826	4	US-09-658-824-330
21	495	30.3	826	4	US-09-959-873B-18
22	495	30.3	826	4	US-09-949-016-6089
23	495	30.3	826	4	US-09-967-388-4
24	495	30.3	826	5	PCT-US96-10251-2
25	489	29.9	810	1	US-08-785-241-7
26	477.5	29.2	875	1	US-08-785-241-5
27	466	28.5	485	3	US-09-374-454-4

28	466	28.5	870	1	US-08-785-241-4	Sequence 4, Appli
29	466	28.5	870	2	US-09-374-454-6	Sequence 6, Appli
30	444	27.1	373	3	US-08-480-473B-3	Sequence 3, Appli
31	444	27.1	373	3	US-08-915-213-3	Sequence 3, Appli
32	444	27.1	373	3	US-09-235-217-3	Sequence 3, Appli
33	444	27.1	373	5	PCT-US96-10251-3	Sequence 3, Appli
34	444	27.1	805	2	US-08-480-473B-4	Sequence 4, Appli
35	444	27.1	805	3	US-08-915-213-4	Sequence 4, Appli
36	444	27.1	805	3	US-09-235-217-4	Sequence 4, Appli
37	444	27.1	805	5	PCT-US96-10251-4	Sequence 4, Appli
38	345.5	21.1	613	4	US-09-438-833-6	Sequence 6, Appli
39	323	19.7	756	4	US-09-438-833-11	Sequence 11, Appli
40	297	18.2	595	4	US-09-949-016-7205	Sequence 7205, Ap
41	294	18.0	590	2	US-08-785-310A-5	Sequence 5, Appli
42	293.5	17.9	594	2	US-08-785-310A-6	Sequence 6, Appli
43	281	17.2	1507	4	US-09-914-259-37	Sequence 37, Appli
44	263	16.1	238	4	US-09-949-016-9161	Sequence 9161, Ap
45	256	15.6	570	4	US-09-923-684-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-09-919-039-149
; Sequence 149, Application US/09919039
; Patent No. 6727066
; GENERAL INFORMATION:
; APPLICANT: Kaefer, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919, 039
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 149
; LENGTH: 827
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6727066 1250434CD1
US-09-919-039-149

Query Match 30.4%; Score 498; DB 4; Length 827;
Best Local Similarity 46.2%; Pred. No. 1,4e-45;
Matches 115; Conservative 26; Mismatches 60; Indels 48; Gaps 6;
QY 9 RSNTELRKESRDPAARSRSQETEVLYOLAHTLPFARGVSAHLDKASIMRLTTSYLRMHR 68
13 RISSRRKESKRPAARSRSKSESVFVELAHQPLPHNVSSHLDKASVMLRTTSYLRVRK 72
Db 69 LCAAGKRGKRGATRLPEGPGFRHGRHGLPVKCOQA-----PGPOSVD 118
73 LIDAG-----DLIDBDMKQNMCFYLKALDGFVMVL 104
QY 119 LCSSSLIH-----NPTPG-TNFSLELIGHSIPFPIHPDOEELQDALTPPNTSKKLEA 172
Db 105 TDDGDMYISDNNVKWGLTQF--ELTGHSVFPTTHCDHEBRKEMLTTHNGVLVKKKEQ 162
QY 173 PTERHPSLRKSTLTSGRTLNKAATWVKLHSGHMAVAKYPPAQSPPASPSPPLOC 232
Db 163 NTPRSFPLRKCTLTSGRTNKSATWVKLCTGHVHTDTSNPOQC--YKKEPMTIC 220
QY 233 LVLTICAIIP 241
Db 221 LVLTICPIP 229
RESULT 2
US-09-438-833-3

```

? Sequence 3 Application US/0943833
? Patent No. 6436654
? GENERAL INFORMATION:
? APPLICANT: Pharmacia & Upjohn
? TITLE OF INVENTION: Protein variants
? FILE REFERENCE: 1848
? CURRENT APPLICATION NUMBER: US/09/438.833
? CURRENT FILING DATE: 1999-11-12
? NUMBER OF SEQ ID NOS: 15
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 3
? LENGTH: 245
? TYPE: PRT
? ORGANISM: Artificial Sequence
? FEATURE:
? OTHER INFORMATION: Description of Artificial Sequence: Subdomain
? OTHER INFORMATION: 1-245 of human HIF-1 alpha
? US-09-438-833-3

```

[illegible]

```

RESULT 3
US-09-438-833-4
; Sequence 4, Application US/09438833
; Patent No. 6436654
; GENERAL INFORMATION:
; APPLICANT: Pharmacia & Upjohn
; TITLE OF INVENTION: Protein variants
; FILE REFERENCE: 1848
; CURRENT APPLICATION NUMBER: US/09/438.833
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Subdomain
; US-09-438-833-4

```

Query March 30.3% Score 495; DB 4; Length 330;
Best Local Similarity 46.2%; Pred. No. 7, 6e-46;
Matches 114; Conservative 26; Mismatches 55; Indels 48; Gaps 6;

Cy 11 NTELRKESRPAASRRSQEETVLYOLAHITPFAFGVSAHLDKDSIMFLTITSYMRRLC 70
::: ::::: :
Db 14 SSESRKESRPAASRRSKSEVEFEALHQLPLPNVSSHLDKDSVMFLTITSYRVRLTL 73

```

QY  AACGKGRGATGRLLEPGCGFRHGRHGRGLGVGCGCOA-----PPQGSVDLC 120
    |||
Db  DAG-----DLIEDMKRQNMCFYLKALDGFVAVLD 105
    |||
QY  121 SSSLIH----NPRQG-TNFSLELIGHSIPDFIHPCDQEBLQDALTERPNLSKKKLEAPT 174
    |||
Db  106 DGDGVITSDNVKKWGLQF--ELTGHGVDFPTPCDHEKREMLTHRNGLVYKKGEQNT 163
    |||
QY  175 ERFSLIRMSSTITSGRGLNLKAAWVKVLHSGHMRAYKPPAQISPGSRSRSEPPQLCV 234
    |||
Db  164 QRSFFLRMCITLTSRGRTWNIKSATWVKVLCGHIHYVDITNSNQPGG--YKRPMTCLV 221
    |||
QY  235 LICEALP 241
    |||
Db  222 LICEPLP 228

```

```

RESULT 4
US-09-967-388-2
; Sequence 2, Application US/09967388
; Patent No. 6838430
;
GENERAL INFORMATION:
;
APPLICANT: JEFFEREY M. ARBETT
;
TITLE OF INVENTION: USE OF HIF-1ALPHA VARIANTS TO ACCELERATE
;
FILE REFERENCE: UC077.001A
;
CURRENT APPLICATION NUMBER: US/09/967,388
;
CURRENT FILING DATE: 2001-09-28
;
NUMBER OF SEQ ID NOS: 4
;
SOFTWARE: FastSeq for Windows Version 4.0
;
SEQ ID NO 2
;
LENGTH: 623
;
TYPE: PRT
;
ORGANISM: HUMAN
;
US-09-967-388-2

```

[illegible]

```

RESULT 5
US-09-438-833-5
; Sequence 5, Application US/09438833
; Patent No. 6438654
; GENERAL INFORMATION:
; APPLICANT: Pharmacia & Upjohn
; TITLE OF INVENTION: Protein variants
; FILE REFERENCE: 1848
; CURRENT APPLICATION NUMBER: US/09/438,833
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1

```



```
RESULT 8
US-08-480-473B-2
; Sequence 2, Application US/08480473B
; Patent No. 5882914
; GENERAL INFORMATION:
; APPLICANT: Semenza, Gregg L.
; TITLE OF INVENTION: HYPOXIA INDUCIBLE FACTOR-1 AND METHOD OF USE
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,473B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/053001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 826 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-480-473B-2

Query Match          30.3%; Score 495; DB 2; Length 826;
Best Local Similarity 46.2%; Pred. No. 3e-45;
Matches 114; Conservative 26; Mismatches 59; Indels 48; Gaps 6;

QY 11 NTELRKESRDARSRSQETEVLYQLAHTLPFARGVSAHLDKASIMRLTISYLRMRLC 70
   :::::::::::::::::::::
DB 14 SSRRRKEKSRDARSRSKSESEVFYEIAHQLPHPNVSSHLDKASVRLTISYLRVAKL 73
QY 71 AAGGKGRAIGRLLPBGPGFRGHTHRGRHGLFVGKCGQA-----PGQSYDLC 120
   |||
DB 74 DAG-----DLDIEDMKQNMCFYKALDGFVWVLTLD 105
QY 121 SSSLIH-----NPTPG-TNFSLELIGHISIPDFIHPDQOEIADALTPRPYLSKKLEAPT 174
   :::::::::::::::::::::
DB 106 DGDMIYISDNVNMKMGITQF--ELTGHISVDFTHPCDHEMRMLTHRNLGVKKGEQNT 163
QY 175 ERHPSLRMKSTLTISRGRTLNKAAWTWVYLHCSGMRAYKPPAQTSPAGSPRSEPPLOCLV 234
   :::::::::::::::::::::
DB 164 QRSFFLRMKCTLTISRGTNMIKSAWTWVYLHCTGHIHYVDYNSNPQCG--YKKPMTCLV 221
QY 235 LICEAIP 241
   |||
DB 222 LICEPIP 228

RESULT 9
US-08-915-213-2
; Sequence 2, Application US/08915213
; Patent No. 6020462
; GENERAL INFORMATION:
; APPLICANT: Semenza, Gregg L.
; TITLE OF INVENTION: HYPOXIA INDUCIBLE FACTOR-1 AND METHOD OF USE
; NUMBER OF SEQUENCES: 64
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/915,213
; FILING DATE: 20-AUG-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/480,473
; FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/053001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 826 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-915-213-2

Query Match          30.3%; Score 495; DB 3; Length 826;
Best Local Similarity 46.2%; Pred. No. 3e-45;
Matches 114; Conservative 26; Mismatches 59; Indels 48; Gaps 6;

QY 11 NTELRKESRDARSRSQETEVLYQLAHTLPFARGVSAHLDKASIMRLTISYLRMRLC 70
   :::::::::::::::::::::
DB 14 SSRRRKEKSRDARSRSKSESEVFYEIAHQLPHPNVSSHLDKASVRLTISYLRVAKL 73
QY 71 AAGGKGRAIGRLLPBGPGFRGHTHRGRHGLFVGKCGQA-----PGQSYDLC 120
   |||
DB 74 DAG-----DLDIEDMKQNMCFYKALDGFVWVLTLD 105
QY 121 SSSLIH-----NPTPG-TNFSLELIGHISIPDFIHPDQOEIADALTPRPYLSKKLEAPT 174
   :::::::::::::::::::::
DB 106 DGDMIYISDNVNMKMGITQF--ELTGHISVDFTHPCDHEMRMLTHRNLGVKKGEQNT 163
QY 175 ERHPSLRMKSTLTISRGRTLNKAAWTWVYLHCSGMRAYKPPAQTSPAGSPRSEPPLOCLV 234
   :::::::::::::::::::::
DB 164 QRSFFLRMKCTLTISRGTNMIKSAWTWVYLHCTGHIHYVDYNSNPQCG--YKKPMTCLV 221
QY 235 LICEAIP 241
   |||
DB 222 LICEPIP 228

RESULT 10
US-09-148-547-2
; Sequence 2, Application US/09148547
; Patent No. 6124131
; GENERAL INFORMATION:
; APPLICANT: Semenza, Gregg L.
; TITLE OF INVENTION: Hypoxia Inducible Factor-1 and Methods of Use
; FILER REFERENCE: 07265/151001
; CURRENT APPLICATION NUMBER: US/09/148,547
; CURRENT FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 826
```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-148-547-2

```

Query Match	30.3%	Score 495;	DB 3;	Length 826;
Best Local Similarity	46.2%;	Pred. No. 3e-45;		
Matches 114;	Conservative 26;	Mismatches 59;	Indels 48;	Gaps 6;

[illegible]

RESULT 11
US-09-235-217-2
; Sequence 2, Application US/09235217

Query Match	30.3%;	Score 495;	DB 3;	length 826;
Best Local Similarity	46.2%;	Pred. No. 3e-45;		

Matches 114; Conservative 26; Mismatches 59; Indels 48; Gaps 6;

```

Qy 11 NTLRREKREKRDAAARSROQETEVLTOLAHATLPPFAGVSAHLIDKASIMRLTISYLRMRCLC 70
Db 14 SSRRREKREKRDAAARSRSKSESEVFPYELAHQPLPLPHNVAASHLDKASVYRLTISYLRVAKKL 73
Qy 71 AAGKRGKRGATGKLLPBGPGFGRHGRHGRHGLPRVKKQQA-----PGRQSDVLC 120
Db 74 DAG-----DLIDEDDKAKNNCFYKALDGFVWVLTLD 105
Qy 121 SSSLIH-----NPTPG-TWPSLELIGHSPDFIHPQOEBLQDALTPRPNLSSKKLEBAPT 174
Db 106 DGMITISDNVKNWYKGLTGF--ELVGHSEVPFTPHCOHEBRERMLTTRHNGVYKKGXEQNT 163
Qy 175 ERHFSLSRMKSTLTSRGRTLNKLAATWKLKJLSCGMRKRAYKPAQOTSPPGSPSESEPLOCIV 234
Db 164 QRSFFLRMKCTLTISRGTNNIKSATWKLHCTGHVHDVITNSNQPOCG--YKRPMTCLV 221
Qy 235 LICEAIP 241
Db 222 LICEPIP 228

```

RESULT 12
US-09-380-662-23
; Sequence 23, Application US/09380662

Query Match	30.3%;	Score 495;	DB 3;	Length 826;
Best Local Similarity	46.2%;	Pred. No. 3e-45;		
Matches 114;	Conservative 26;	Mismatches 59;	Indels 48;	Gaps 6;

QY	1	NTFLRREKSRDAAARSROETVYQLAHNTLPFARGVSAHLDDKASIMRLTISYLRMRILC	70
Db	14	SSRRREKSRDAAARSRSKSESVFELAHQLPFARHNVSHLDDKASVNRLLTISYLRAKLL	73
QY	71	AAGKKRGATGRLLEPGGFRHGTTHRRGRHLEPVGKQQA-----PGQSVILC	120
Db	74	DAG-----DLIDEDMKQNMCFYLKALDDGVWVLTLD	105
QY	121	SSSLIH-----NPTRG-TNFSLELTGHSIPDPIPCPOEQLDLPTRPMLSKKLEAPT	174
Db	106	DGDMLTISNNVKKWGLTQF--ELTGHSVFDFTHPCHEERREMLTHRNGILVKKGEQNT	163
QY	175	ERHFSLRMKSITLTSRGRTLLKAAATWKLSCSGHMRAYKPPACTSPASPRSEPPLOCLV	234
Db	164	QRSFFLRMKCTLTSGRMTNIKSAIWKVYLCTGHIHYVDNNSNQPGQG--YKDPMTCLV	221
QY	235	LICEAP	241
Db	222	LICEPI	228

RESULT 13

Db	164	Q R S F P L M K C T L T S R G R T M N I K S A T W K V L H C T G H I H V Y D I N S N O P Q G - - Y K P P M T C L V	221
Qy	235	L I C E A I P	241
Db	222	L I C E P I P	228

Search completed: June 15, 2005, 16:01:42
Job time : 43 secs

This Page Blank (copy)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 15:36:58 ; Search time 162 Seconds
(without alignments)
732.935 Million cell updates/sec

Title: US-09-896-791B-3

Perfect score: 1636
Sequence: 1 MALGLQRVRSNTBLRKEKSR.....TESSLPSWVLMALNRKNCPG 307

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

1: geneseqp1980s:*\n2: geneseqp1990s:*\n3: geneseqp2000s:*\n4: geneseqp2001s:*\n5: geneseqp2002s:*\n6: geneseqp2003as:*\n7: geneseqp2003bs:*\n8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1636	100.0	307	5	AAU75902 Mouse inh
2	1636	100.0	307	6	ABR62227 Human inh
3	1278	78.1	324	6	ABR83343 Human NOV
4	855.5	52.3	662	2	AAV06295 Mouse tra
5	831	50.8	632	4	AAH93326 Human pro
6	803	49.1	790	6	AAO16439 Human nuc
7	802	49.0	667	5	AAE24222 Human nuc
8	797.5	48.7	747	6	AAO16417 Human HIF
9	797	48.6	407	5	ABG66737 Human nov
10	795	48.6	705	6	AAE34826 Human hyp
11	573	35.0	648	4	AAH94934 Human pro
12	518.5	31.7	199	4	AAH93710 Human pro
13	498	30.4	513	8	AD116229 Human nuc
14	498	30.4	827	7	ADE25733 Human pro
15	498	30.4	827	8	ADE76984 Human pro
16	497	30.4	825	7	ADDA4855 Rat Prote
17	495	30.3	245	3	AAV94627 HIF-1alp
18	495	30.3	330	3	AAV94628 HIF-1alp
19	495	30.3	623	6	ABR40379 Human HIF
20	495	30.3	642	3	AAV84168 A variant
21	495	30.3	652	3	AAV94629 HIF-1alp
22	495	30.3	669	3	AAV84167 A variant
23	495	30.3	697	3	AAV84166 A variant
24	495	30.3	701	3	AAV84173 A variant
25	495	30.3	710	3	AAV84172 A variant

26	495	30.3	724	3	AAV84171 A variant
27	495	30.3	735	6	ABR82375 Hypoxia-1
28	495	30.3	735	8	ADN75066 Human hyp
29	495	30.3	749	3	AAV84170 A variant
30	495	30.3	789	3	AAV84169 A variant
31	495	30.3	789	6	ADA18535 Human hyp
32	495	30.3	789	6	ADA18533 Human hyp
33	495	30.3	789	6	ADA18534 Human hyp
34	495	30.3	813	4	ADA18533 Human hyp
35	495	30.3	826	2	AAV06557 Human hyp
36	495	30.3	826	2	AAH80418 amino aci
37	495	30.3	826	2	AAV06289 Human tra
38	495	30.3	826	3	AAV69407 A wild ty
39	495	30.3	826	3	AAV94640 Human hyp
40	495	30.3	826	4	AAV76854 Human lun
41	495	30.3	826	5	AAV85509 Clone #19
42	495	30.3	826	5	AAU77620 Human hyp
43	495	30.3	826	5	AAU77602 Human hyp
44	495	30.3	826	5	AAU77619 Human hyp
45	495	30.3	826	5	AAU77607 Human hyp

ALIGNMENTS

RESULT 1	AAU75902 standard; protein; 307 AA.
ID	AAU75902
XX	AAU75902;
AC	08-MAY-2002 (first entry)
DT	08-MAY-2002 (first entry)
XX	Mouse inhibitory PAS domain (IPAS) protein.
DE	Inhibitory PAS domain protein; IPAS; cardiant; vasotropic;
XX	cardiovascular; cerebroprotective; ophthalmological; HIF-1alpha;
KW	hypoxia-induced factor 1alpha; vascular endothelial growth factor;
KW	angiogenesis; ischemic cardiovascular lesion; stroke;
KW	diabetic microvascular disease; tumour; mouse.
XX	Mus musculus.
OS	WO200202609-A1.
XX	10-JAN-2002.
PN	19-JUN-2001; 2001WO-SE001387.
XX	06-JUL-2000; 2000SE-00002551.
PR	(BIOV-) BIOVITRUM AB.
PA	Berkenstam A, Bertilsson G, Poellinger L;
PI	WPI; 2002-164523/21.
XX	N-PSDB; ABK14502.
DR	New nucleic acid encoding inhibitory PAS domain protein, useful for
PT	identifying specific inhibitors for treating e.g. angiogenesis or tumor
XX	growth.
XX	Claim 3; Fig 1; 44pp; English.
PS	The invention describes an isolated nucleic acid encoding the
CC	biologically active inhibitory PAS domain protein or its functionally
CC	equivalent modifications. IPAS forms a non-functional heterodimeric
CC	complex with HIF-1alpha (hypoxia-induced factor 1alpha), impairing
CC	interaction between HIF-1alpha and hypoxia-response elements in genes,
CC	e.g. the gene for vascular endothelial growth factor, so contributes to
CC	control of hypoxic signalling. The nucleic acid and its encoded
CC	polypeptides, are used to identify agents that activate expression of the
CC	gene or stimulate activity of the protein. These agents are useful for

CC inhibiting angiogenesis, particularly where associated with ischaemic
 CC cardiovascular lesions, stroke or diabetic microvascular diseases, and
 CC tumour growth. This is the amino acid sequence of the mouse inhibitory
 CC PAS domain protein (IPAS), described in the method of the invention
 XX

SO Sequence 307 AA;

Query Match 100.0%; Score 1636; DB 5; Length 307;
 Best Local Similarity 100.0%; Pred. No. 1.9e-160;
 Matches 307; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALGLQVRVSNTELRKESRDARSRSQETEVLYQLAHTLPFARGVSAHLDDKASIMRLT 60
 DB 1 MALGLQVRVSNTELRKESRDARSRSQETEVLYQLAHTLPFARGVSAHLDDKASIMRLT 60
 QY 61 ISYLRMRLCAAGKGRATGRLLPEPGGFRHGTTHRGHGLPVGCGQAPGQSVDLIC 120
 DB 61 ISYLRMRLCAAGKGRATGRLLPEPGGFRHGTTHRGHGLPVGCGQAPGQSVDLIC 120
 QY 121 SSSLINHPPTGNTFSLFELIGHSTFDPIHPCDQBELQDALTPRPNLSSKKLEAPTEHFSL 180
 DB 121 SSSLINHPPTGNTFSLFELIGHSTFDPIHPCDQBELQDALTPRPNLSSKKLEAPTEHFSL 180
 QY 181 RMKSTLSRGTNLKATWVKVHSCGHRAYRPAQTSRPSGSRSEPPLOCLVLCIAI 240
 DB 181 RMKSTLSRGTNLKATWVKVHSCGHRAYRPAQTSRPSGSRSEPPLOCLVLCIAI 240
 QY 241 POLPFHDGATLGLPOEKTPISTLFTPLMKALLCLVWRMPVOVLQKGTSSLPSPVLMAL 300
 DB 241 POLPFHDGATLGLPOEKTPISTLFTPLMKALLCLVWRMPVOVLQKGTSSLPSPVLMAL 300
 QY 301 NRKNCPPG 307
 DB 301 NRKNCPPG 307

RESULT 2

ABR62227 standard; protein; 307 AA.

ABR62227;

08-SEP-2003 (first entry)

Mouse inhibitory PAS (Per/Arnt/Sim) domain protein IPAS.

Mouse; IPAS; inhibitory PAS domain protein; hypoxia; vasotropic;

cardiac; vulnerable; cerebroprotective; antiulcer; gene therapy.

Mus sp.

Key Location/Qualifiers

Region 14..25 /label= Basic

Region 26..41 /label= Helix

Region 42..55 /label= Loop

Region 56..67 /label= Helix

Domain 103..170 /label= PAS A

Domain 235..307 /label= PAS B

WO2003045440-A1.

05-JUN-2003.

28-NOV-2002; 2002WO-SE002198.

28-NOV-2001; 2001US-0333513P.

PA (ANGI-) ANGIOGENETICS SWEDEN AB.
 XX Makino Y, Cao Y, Poellinger L;
 XX WPI; 2003-505168/47.
 DR N-PDB; ACC83867.

PT New pharmaceutical composition comprising IPAS, useful for increasing
 PT angiogenesis, or for treating a hypoxia-mediated condition in a cell,
 PT group of cells or organisms under hypoxic conditions, e.g. stroke or
 PT coronary heart disease.

PS Claim 4; Fig 1; 45pp; English.

CC The present sequence is the protein sequence of a novel murine basic
 CC helix-loop-helix/PAS (Per/Arnt/Sim) protein, termed inhibitory PAS domain
 CC protein or IPAS, that is related to hypoxia-inducible transcription
 CC factors (HIFs). IPAS has no transactivation function and demonstrates
 CC dominant negative regulation of HIF-mediated control of gene expression.
 CC Ectopic expression of IPAS in hepatoma cells selectively impaired
 CC induction of genes involved in adaptation to a hypoxic environment, and
 CC also resulted in retarded tumour growth and reduced tumour vascular
 CC density in vivo. In mice, IPAS was predominantly expressed in Purkinje
 CC cells and in corneal epithelium. Application of an IPAS antisense
 CC oligonucleotide to the mouse cornea induced angiogenesis under normoxic
 CC conditions, and unmasked hypoxia-dependent induction of vascular
 CC endothelial growth factor gene expression in hypoxic cornea cells,
 CC indicating a novel regulation of angiogenesis and maintenance of an
 CC avascular phenotype. Claimed pharmaceutical compositions comprising IPAS
 CC antisense sequences can be used to treat increased angiogenesis, to
 CC stimulate HIF-alpha function, to treat hypoxia-related conditions such as
 CC ischaemia, coronary heart disease, wound healing, stroke or diabetic
 CC ulceration, and to maintain normal cell functions under hypoxia
 XX

SO Sequence 307 AA;

Query Match 100.0%; Score 1636; DB 6; Length 307;
 Best Local Similarity 100.0%; Pred. No. 1.9e-160;
 Matches 307; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALGLQVRVSNTELRKESRDARSRSQETEVLYQLAHTLPFARGVSAHLDDKASIMRLT 60
 DB 1 MALGLQVRVSNTELRKESRDARSRSQETEVLYQLAHTLPFARGVSAHLDDKASIMRLT 60
 QY 61 ISYLRMRLCAAGKGRATGRLLPEPGGFRHGTTHRGHGLPVGCGQAPGQSVDLIC 120
 DB 61 ISYLRMRLCAAGKGRATGRLLPEPGGFRHGTTHRGHGLPVGCGQAPGQSVDLIC 120
 QY 121 SSSLINHPPTGNTFSLFELIGHSTFDPIHPCDQBELQDALTPRPNLSSKKLEAPTEHFSL 180
 DB 121 SSSLINHPPTGNTFSLFELIGHSTFDPIHPCDQBELQDALTPRPNLSSKKLEAPTEHFSL 180
 QY 181 RMKSTLSRGTNLKATWVKVHSCGHRAYRPAQTSRPSGSRSEPPLOCLVLCIAI 240
 DB 181 RMKSTLSRGTNLKATWVKVHSCGHRAYRPAQTSRPSGSRSEPPLOCLVLCIAI 240
 QY 241 POLPFHDGATLGLPOEKTPISTLFTPLMKALLCLVWRMPVOVLQKGTSSLPSPVLMAL 300
 DB 241 POLPFHDGATLGLPOEKTPISTLFTPLMKALLCLVWRMPVOVLQKGTSSLPSPVLMAL 300
 QY 301 NRKNCPPG 307
 DB 301 NRKNCPPG 307

RESULT 3

ABR83343 standard; protein; 324 AA.

ABR83343;

06-OCT-2003 (first entry)

DE Human NOV9a protein SEQ ID NO:20.
 XX
 KM Human; NOVX; cytosolic; antidiabetic; neuroprotective; antiparkinsonian;
 KM anorectic; gene therapy; vaccine; cancer; neurodegenerative disorder;
 KM Parkinson's disease; metabolic disorder; diabetes; obesity;
 KM tissue typing.
 XX
 OS Homo sapiens.
 XX
 PN WO2003052061-A2.
 XX
 PD 26-JUN-2003.
 XX
 PF 03-DEC-2002; 2002WO-US038821.
 XX
 PR 17-DEC-2001; 2001US-0341477P.
 PR 17-DEC-2001; 2001US-0341540P.
 PR 20-DEC-2001; 2001US-0342552P.
 PR 31-DEC-2001; 2001US-0344903P.
 PR 17-APR-2002; 2002US-0373288P.
 PR 15-MAY-2002; 2002US-0380981P.
 PR 17-MAY-2002; 2002US-0381495P.
 PR 28-MAY-2002; 2002US-0383744P.
 PR 29-MAY-2002; 2002US-0384024P.
 PR 07-AUG-2002; 2002US-0401788P.
 PR 26-AUG-2002; 2002US-0406353P.
 PR 31-OCT-2002; 2002US-0422756P.
 PR 02-DEC-2002; 2002US-00307928.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Alsobrook JP, Anderson DW, Boldog FL, Burgess CE, Catterton E;
 PI Edinger SR, Gorman L, Guo X, Ji W, Kekuda R, Li L, Paturajan M;
 PI Rieger DK, Shenoy SG, Spytek KA, Vernet CAM, Voss EZ, Zhong M;
 DR WPI, 2003-533005/50.
 DR N-PSDB; ACP06242.
 XX
 PT New NOVX polypeptide, useful for preparing a composition for treating or
 PT preventing e.g. cancer, neurodegenerative disorders such as Parkinson's
 PT disease, or metabolic disorders such as diabetes or obesity, or for
 PT tissue typing.
 XX
 PS Claim 1; Page 130; 190pp; English.
 XX
 CC ACP06242 encode the human NOVX proteins given in ABR83334 to
 CC ABR83343, designated NOV1a, NOV2a, NOV3a, NOV4a, NOV5a, NOV6a,
 CC NOV7a, NOV8a and NOV9a respectively. NOVX sequences can have cytosolic,
 CC antidiabetic, neuroprotective, antiparkinsonian and anorectic activities,
 CC and can be used in vaccines and gene therapy. The NOVX polypeptides can
 CC be used for preparing a composition for treating or preventing a
 CC pathology associated with the NOVX-polypeptides e.g. cancer,
 CC neurodegenerative disorders such as Parkinson's disease, or metabolic
 CC disorders such as diabetes or obesity, or for tissue typing
 CC
 XX
 SQ Sequence 324 AA;
 Query Match 78.1%; Score 1278; DB 6; Length 324;
 Best Local Similarity 79.9%; Pred. No. 2.5e-123;
 Matches 250; Conservative 14; Mismatches 41; Indels 8; Gaps 3;
 QY 1 MALGLOFVRSNTELRKESRDAAARSRSQETEVLYQLAHTLPFARGVSAHLDKASIMRLT 60
 DB 1 MALGLOFARSTTELRKESRDAAARSRSQETEVLYQLAHTLPFARGVSAHLDKASIMRLT 60
 QY 61 ISYIRHMRLLCAAGSGRRATGRLLPBGPGFRHGTTHRRGRGLFVGCCQAPEQSYDLC 120
 DB 61 ISYIRHMRLLCAAGSGRRATGRLLPBGPGFRHGTTHRRGRGLFVGCCQAPEQSYDLC 120
 QY 121 SSSLINHPPTGNTSLSLIGHSIFDPFIHPCDQBELQDALPFRPVLSKKKL EAPTERHFSL 180
 DB 121 SSSLINHPPTGNTSLSLIGHSIFDPFIHPCDQBELQDALPFRPVLSKKKL EAPTERHFSL 180
 DB 121 SPSLPRIPTPTGNTSLSLIGHSIFDPFIHPCDQBELQDALPFRPVLSKKKL EAPTERHFSL 180

QY 181 RMKSTLTSRGRTINLKAATKVLHCGHMRAYYPAPQTSPPAGSPRSEPPQLCVLICEAT 240
 DB 181 RMKSTLTSRGRTINLKAATKVLHCGHMRAYYPAPQTSPPAGSPRSEPPQLCVLICEAT 240
 QY 241 POLPFDHGAFLGAPQETPTISTPTPLMKL-----LCVKKRPVQV-LOGKGTSSLPSS 294
 DB 241 PSFLSTGATLGLPQEA--SPLLSPYPCVGRHSCLCFVREWPDIQAGMGSDPSLLPL 298
 QY 295 WVLWALNRKNCPG 307
 DB 299 EVMWALNGNCAG 311

RESULT 4
 AAY06295
 ID AAY06295 standard; protein; 662 AA.
 AC AAY06295;
 DT 23-AUG-1999 (first entry)
 XX
 DE Mouse transcription regulator MOP7.
 XX
 KM MOP7, member of the PAS superfamily; bHLH-PAS; mouse;
 KM transcription regulator; hypoxia inducible factor 3 alpha.
 OS Mus musculus.
 XX
 PN W09928464-A2.
 XX
 PD 10-JUN-1999.
 XX
 PF 27-NOV-1998; 98MO-US025314.
 PR 28-NOV-1997; 97US-0066863P.
 PA (WISC) WISCONSIN ALUMNI RES FOUND.
 XX
 PI Bradford CA, Gu YZ, Hogenesch JB;
 DR WPI, 1999-371120/31.
 DR N-PSDB; AAX58986.
 XX
 PT Developmental signal transduction associated proteins.
 XX
 PS Claim 6; Page 101; 106pp; English.
 XX
 CC The present sequence represents mouse MOP7, a novel member of the PAS
 CC superfamily, where PAS stands for PER/ARNT/SIM domains. MOP7 cDNA (see
 CC AAX58986) was identified in a search of murine ESTs designed to identify
 CC bHLH-PAS proteins, and by RACE amplification of lung cDNA. MOP7 was
 CC characterised as hypoxia inducible factor 3 alpha (HIF 3 alpha). Its
 CC expression profile is distinct from that of HIF 1 alpha (see AAY06289),
 CC HIF 2 alpha (see AAY06290), MOP3 (see AAY06291), Ah receptor and Ah
 CC receptor nuclear translocator (ARNT), suggesting a different functional
 CC role. MOP7 probably regulates the same genes as HIF 1 alpha and 2 alpha,
 CC as evidenced by its dimerisation with the same partners (ARNT, MOP3) and
 CC recognition of the same core response element. MOP7 may have a functional
 CC role associated with response to low oxygen in the tissues in which it is
 CC expressed. The invention provides novel MOPs 2-9 nucleic acids (see
 CC AAX58981-88) and proteins (see AAY06289-97). These are useful in a
 CC variety of research, diagnostic and therapeutic applications. Several of
 CC the MOPs are alpha-class hypoxia-inducible factors. Others are involved
 CC in circadian signal transduction
 CC
 XX
 SQ Sequence 662 AA;
 Query Match 52.3%; Score 855.5; DB 2; Length 662;
 Best Local Similarity 68.4%; Pred. No. 4.2e-79;
 Matches 182; Conservative 5; Mismatches 24; Indels 55; Gaps 4;
 QY 9 RSTELRKESRDAAARSRSQETEVLYQLAHTLPFARGVSAHLDKASIMRLTISYLRMR 68

Db 7 RSNTLEKESRDARSRSQETEVLYQLAHTLPFARGVSAHLDKASIMRLTISYLEMR 66
 Qy 69 LCAAGKRGKRGATGRLLEPGGPFHGTTRGRHGLPVGKCCQAPGPOSDLC----- 120
 Db 67 LCAAGENN-----QVEKGSPDLACTYKALEGF 94
 Qy 121 -----SSSLIHNPTPGTNFS-LELIGHSTIDFPHPCDQELQDALTPRPVLSKKL 170
 Db 95 VMVLTAGDMAVYLEENYSKILGSLQLELIGHSTIDFPHPCDQELQDALTPRPVLSKKL 154
 Qy 171 EAPTERHPSLRMKSSTLTSRGRTLNKKAATWVVLHCSGHRMAYKRPAPQTSPPAGSPRSEPL 230
 Db 155 EAPTERHPSLRMKSSTLTSRGRTLNKKAATWVVLHCSGHRMAYKRPAPQTSPPAGSPRSEPL 214
 Qy 231 QCVLVICCAIP-----QLPFHDGATL 251
 Db 215 QCVLVICCAIPHPASLEPPLGRGAFL 240
 Db
 RESULT 5
 AAB93326
 ID AAB93326 standard; protein; 632 AA.
 XX
 AC AAB93326;
 DT 26-JUN-2001 (first entry)
 DE Human protein sequence SEQ ID NO:12422.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN EPI074617-A2.
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-00116126.
 XX
 PR 29-JUL-1999; 99JP-00248036.
 PR 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.
 PT
 PS Claim 8; SEQ ID NO 12422; 2537bp + Sequence listing; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesizing polynucleotides.

CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH1633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention
 CC
 XX
 SQ Sequence 632 AA;
 Query Match 50.8%; Score 831; DB 4; Length 632;
 Best Local Similarity 68.2%; Pred. No. 1.4e-76;
 Matches 178; Conservative 11; Mismatches 30; Indels 42; Gaps 5;
 Qy 1 MALGQVRSNTLEKESRDARSRSQETEVLYQLAHTLPFARGVSAHLDKASIMRLT 60
 Db 1 MALGQVRSNTLEKESRDARSRSQETEVLYQLAHTLPFARGVSAHLDKASIMRLT 60
 Qy 61 ISYLRMRLCAAGKRGKRGATGRLLEPGGPFHGTTRGRHGLPVGKQ----- 109
 Db 61 ISYLRMRLCAAG-----EMNQVAGGEPDLACTYKALEGFVMVL 100
 Qy 110 QARGPOSDLCSSSLIHNPTPGTNFS-LELIGHSTIDFPHPCDQELQDALTPRPVLSKK 168
 Db 101 TARGDMAY-----LSENVSKILGSLQLELIGHSTIDFPHPCDQELQDALTPRPVLSRR 154
 Qy 169 KLEAPTERHPSLRMKSSTLTSRGRTLNKKAATWVVLHCSGHRMAYKRPAPQTSPPAGSPRSE 228
 Db 155 KLEAPTERHPSLRMKSSTLTSRGRTLNKKAATWVVLHCSGHRMAYKRPAPQTSPPAGSPRSE 214
 Qy 229 PLQCVLVICCAIPQLPFHDGATL 249
 Db 215 PLQCVLVICCAIP-----HPGS 231
 Db
 RESULT 6
 AA016439
 ID AA016439 standard; protein; 790 AA.
 XX
 AC AA016439;
 DT 10-APR-2003 (first entry)
 DE Human nucleic acid-associated protein (NAAP) - SEQ ID NO 36.
 XX
 KW Human; nucleic acid-associated protein; NAAP; arteriosclerosis;
 KW cell proliferative disorder; atherosclerosis; cirrhosis; hepatitis; AIDS;
 KW cancer; developmental disorder; renal tubular acidosis; anaemia; asthma;
 KW mental retardation; neurological disorder; Alzheimer's disease; epilepsy;
 KW Parkinson's disease; autoimmune disorder; inflammatory disorder; allergy;
 KW Crohn's disease; transgenic animal; animal model.
 XX
 OS Homo sapiens.
 XX
 PN WC02003000864-A2.
 PD 03-JAN-2003.
 XX
 PF 20-JUN-2002; 2002WO-US021179.
 XX
 PR 22-JUN-2001; 2001US-0300518P.
 PR 29-JUN-2001; 2001US-0301787P.
 PR 29-JUN-2001; 2001US-0301792P.
 PR 29-JUN-2001; 2001US-0301892P.
 PR 29-JUN-2001; 2001US-0301893P.
 PR 06-JUL-2001; 2001US-0303405P.
 PR 06-JUL-2001; 2001US-0303442P.
 PR 15-MAR-2002; 2002US-0364438P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Gandhi AR, Swarnakar A, Hafalia AJA, Warren BA, Emerling BM;

PI Arviizu CS, Ison CH, Honchell CD, Lee EA, Yue H, Foreythe IU;
 PI Rankumar J, Griffin JA, Yang J, Sanjanwala MM, Baughn MR;
 PI Borowsky ML, Yao MG, Wala NK, Bandman O, Lal PG, Becha SD, Lee SY;
 PI Richardson TW, Elliott VS, Luo W, Tang YT, Zebardjian Y, Lu Y;
 DR WPI; 2003-201420/19.
 XX N-PSDB; AAL51589.
 XX
 XX New nucleic acid-associated proteins and polynucleotides, useful for
 PT diagnosing, treating or preventing cell proliferative (e.g. cancer),
 PT neurological (e.g. epilepsy or Parkinson's disease), or autoimmune
 PT disorders (e.g. AIDS).
 XX
 XX Claim 1; Page 274-276; 312pp; English.
 XX
 XX The invention comprises the amino acid and coding sequences of human
 CC nucleic acid-associated proteins (NAAP). The DNA and protein sequences of
 CC the invention are useful for diagnosing, treating or preventing disorders
 CC associated with aberrant expression of NAAP, such as: cell proliferative
 CC disorders (e.g. arteriosclerosis, atherosclerosis, cirrhosis, hepatitis
 CC or cancer); developmental disorders (e.g. renal tubular acidosis, anaemia
 CC or mental retardation); neurological disorders (e.g. Alzheimer's disease,
 CC Parkinson's disease or epilepsy); and autoimmune/inflammatory disorders
 CC (e.g. AIDS, allergies, asthma or Crohn's disease). The DNA sequences of
 CC the invention are useful for creating transgenic animals to model human
 CC disease. The present amino acid sequence represents a human nucleic acid-
 CC associated protein of the invention
 XX
 XX Sequence 790 AA:
 SQ
 Query Match 49.1%; Score 803; DB 6; Length 790;
 Best Local Similarity 58.7%; Pred. No. 1.5e-73;
 Matches 179; Conservative 11; Mismatches 29; Indels 86; Gaps 6;
 QY 1 MALGLQVR-----SNTLEIRK 16
 DB 1 MALGIQARPALSGVISPPCAPTRNSHPGCGTASPPAPGWPFSQKPGKRWSTTEIRK 60
 QY 17 EKSRDARRSRROETEVLYQLAHTLPFARGVAHLDKASIMRLTSTYLRMRCAAGGKR 76
 DB 61 EKSRDARRSRROETEVLYQLAHTLPFARGVAHLDKASIMRLTSTYLRMRCAAG-- 117
 QY 77 GRATGRLLPEGPGGFRHGTHRRGRGLPVGKQ-----QAPGQSVLDCSSSLI 125
 DB 118 -----EMNGVAGGEBPLDACVYLKALGFMVWVLAEGDMAY-----LS 154
 QY 126 HNPPGTFNS-LELIGHSTFDFIHCDOBELDALTPPNLSKKKLEAPTERHSLRWKS 184
 DB 155 ENVSKHGLSQLELIGHSTFDFIHCDOBELDALTPQOTLSRRKVEAPTERCFSLRWKS 214
 QY 185 TLTSRGTLNKAATWVKVLCGSHMARAYKPPAOTSPPGSPRSEPLQCLVLCIAIPQLP 244
 DB 215 TLTSRGTLNKAATWVKVLCGSHMARAYKPPAOTSPPGSPRSEPLQCLVLCIAIP-- 271
 QY 245 FHDGA 249
 DB 272 -HPGS 275
 XX
 XX RESULT 7
 XX AAE24222
 XX ID AAE24222 standard; protein; 667 AA.
 XX AC AAE24222;
 XX
 XX 23-SEP-2002 (first entry)
 XX
 XX Human HIF-3 protein.
 XX
 XX Human; haematologic malignancy; multidrug resistance; MDR; SUMO-1;
 KW hypoxia inducible factor-1; small ubiquitin-like-modifier; HIF-3;
 KW lymphoid disorder; chronic lymphoproliferative disorder; lymphoma;
 KW myeloid disorder; lymphocytic leukaemia; thrombocythaemia; myeloma;

KW angiogenic myeloid metaplasia; myeloid leukaemia; gene therapy;
 KW polycythaemia vera; hypoxia responsive element; HRE.
 XX
 XX Homo sapiens.
 XX
 XX W0200234291-A2.
 XX
 XX 02-MAY-2002.
 XX
 XX 25-OCT-2001; 2001WO-US049856.
 XX
 XX 26-OCT-2000; 2000US-0243542P.
 XX
 XX (BGM) BRIGHAM & WOMENS HOSPITAL INC.
 XX
 XX Colgan SP;
 XX
 XX WPI; 2002-471427/50.
 XX
 XX N-PSDB; AAD39042.
 XX
 XX Treating a subject (at risk of) having a hematologic malignancy or
 PT multidrug resistance, e.g. lymphoma or myeloma, by administering hypoxia
 PT inducible factor 1 binding molecules or small ubiquitin-like-modifier-1
 PT binding molecules.
 XX
 XX Disclosure; Page 88-91; 92pp; English.
 XX
 XX The invention relates to a method of treating a subject having or at risk
 CC of developing a haematologic malignancy or multidrug resistance (MDR).
 CC The method involves administering hypoxia inducible factor-1 (HIF-1)
 CC binding molecules or small ubiquitin-like-modifier (SUMO)-1 binding
 CC molecules or HIF-1-SUMO-1 complex modulators. mdr1-hypoxia responsive
 CC element (HRE) binding molecules or antisense nucleic acid molecules and
 CC SUMO-1-binding molecules or antisense molecules are useful for treating a
 CC subject having or at risk of developing haematologic malignancy or MDR
 CC (e.g. a lymphoid disorder or a myeloid disorder). The lymphoid disorders
 CC include lymphocytic leukaemia or chronic lymphoproliferative disorders
 CC e.g. lymphoma, myeloma or chronic lymphoid leukaemia. The myeloid
 CC disorders include chronic or acute myeloid leukaemia, e.g. angiodenic
 CC myeloid metaplasia, essential thrombocythaemia or polycythaemia vera. The
 CC invention is used in gene therapy. The present sequence is human HIF-3
 CC protein
 XX
 XX Sequence 667 AA:
 SQ
 Query Match 49.0%; Score 802; DB 5; Length 667;
 Best Local Similarity 68.0%; Pred. No. 1.5e-73;
 Matches 172; Conservative 11; Mismatches 28; Indels 42; Gaps 5;
 QY 9 RSNTELEKESRDAARRSRROETEVLYQLAHTLPFARGVAHLDKASIMRLTSTYLRMR 68
 DB 7 RSTTELEKESRDAARRSRROETEVLYQLAHTLPFARGVAHLDKASIMRLTSTYLRMR 66
 QY 69 LCAAGGKRGRATGRLLPEGPGGFRHGTHRRGRGLPVGKQ-----QAPGQSV 117
 DB 67 LCAAG-----EMNGVAGGEBPLDACVYLKALGFMVWVLAEGDMAY 106
 QY 118 DLGSSSLHNPTGTFNS-LELIGHSTFDFIHCDOBELDALTPPNLSKKKLEAPTER 176
 DB 107 -----LSENVSKHGLSQLELIGHSTFDFIHCDOBELDALTPQOTLSRRKVEAPTER 160
 QY 177 HPSLRKSTLTSRGTLNKAATWVKVLCGSHMARAYKPPAOTSPPGSPRSEPLQCLVLI 236
 DB 161 CFSLRKSTLTSRGTLNKAATWVKVLCGSHMARAYKPPAOTSPPGSPRSEPLQCLVLI 220
 QY 237 CEAIPLPFDGA 249
 DB 221 CEAIPLP-----HPGS 229
 XX
 XX RESULT 8
 XX AAO16417
 XX ID AAO16417 standard; protein; 747 AA.

XX AA016417;
 AC 10-APR-2003 (first entry)
 DT 10-APR-2003 (first entry)
 XX Human nucleic acid-associated protein (NAAP) - SEQ ID NO 14.
 DE Human nucleic acid-associated protein; NAAP; arteriosclerosis;
 XX Human; nucleic acid-associated protein; NAAP; arteriosclerosis;
 XX cell proliferative disorder; atherosclerosis; cirrhosis; hepatitis; AIDS;
 XX cancer; developmental disorder; renal tubular acidosis; anemia; asthma;
 XX mental retardation; neurological disorder; Alzheimer's disease; epilepsy;
 XX Parkinson's disease; autoimmune disorder; inflammatory disorder; allergy;
 XX Crohn's disease; transgenic animal; animal model.
 OS Homo sapiens.
 PN WO200300864-A2.
 XX 03-JAN-2003.
 PD 20-JUN-2002; 2002WO-US021179.
 PF 22-JUN-2001; 2001US-0300518P.
 XX 29-JUN-2001; 2001US-0301787P.
 PR 29-JUN-2001; 2001US-0301792P.
 PR 29-JUN-2001; 2001US-0301892P.
 PR 29-JUN-2001; 2001US-0301893P.
 PR 06-JUL-2001; 2001US-0303405P.
 PR 06-JUL-2001; 2001US-0303442P.
 PR 15-MAR-2002; 2002US-0364438P.
 XX (INCYTE GENOMICS INC.
 PA Gandhi AR, Swarnakar A, Hafalia AUA, Warren BA, Emerling BM,
 XX Aryizu CS, Ison CH, Honchell CD, Lee EA, Yue H, Forsythe IU,
 PI Rankumar J, Griffin JA, Yang J, Sanjanwala MM, Baughn KR;
 PI Borowaky ML, Yao MG, Wajia NK, Bandman O, Lal PG, Becha SD, Lee SY,
 PI Richardson TM, Elliott VS, Luo W, Tang YT, Zebardjian Y, Lu Y;
 XX WPI; 2003-201420/19.
 DR N-PSDB; AAL51567.
 XX New nucleic acid-associated proteins and polynucleotides, useful for
 PT diagnosing, treating or preventing cell proliferative (e.g. cancer),
 PT neurological (e.g. epilepsy or Parkinson's disease), or autoimmune
 PT disorders (e.g. AIDS).
 XX Claim 1; Page 232-234; 312pp; English.
 XX The invention comprises the amino acid and coding sequences of human
 CC nucleic acid-associated proteins (NAAP). The DNA and protein sequences of
 CC the invention are useful for diagnosing, treating or preventing disorders
 CC associated with aberrant expression of NAAP, such as: cell proliferative
 CC disorders (e.g. arteriosclerosis, atherosclerosis, cirrhosis, hepatitis
 CC or cancer); developmental disorders (e.g. renal tubular acidosis, anemia
 CC or mental retardation); neurological disorders (e.g. Alzheimer's disease,
 CC Parkinson's disease or epilepsy); and autoimmune/inflammatory disorders
 CC (e.g. AIDS, allergies, asthma or Crohn's disease). The DNA sequences of
 CC the invention are useful for creating transgenic animals to model human
 CC disease. The present amino acid sequence represents a human nucleic acid-
 CC associated protein of the invention
 XX Sequence 747 AA;
 SQ Query Match 48.7%; Score 797.5; DB 6; Length 747;
 Beat Local Similarity 60.4%; Pred. No. 5.2e-73;
 Matches 180; Conservative 10; Mismatches 35; Indels 73; Gaps 6;
 QY 1 MALGIQVR-----SNTELRK 16
 DB 1 MALGIQVRPALSCGVISPPCAPTRNSHPGCTASPPAPGWPFSORGPGRWSTELRK 60
 QY 17 EKSRDARSRSQETEVLYQLAHTLPFARGVSAHLDRASIMRLTISYLRMRCAAGGR 76

DB 61 EKSRDARSRSQETEVLYQLAHTLPFARGVSAHLDRASIMRLTISYLRMRCAAGGRN 120
 QY 77 GRATGR-----LLEPGCGFGRHGRGRGLPVKGCQAPGPOSVDLCSSSLHNPTGT 132
 DB 121 QVAGENHMMLLKALBGFVMVLTAE-----DM--AVISENVSKHL 160
 QY 133 NFS-LELIGHSTDFPHPCDQELQALTPRPULSKKLEAPFERFSLRMSTLTSSRR 191
 DB 161 GLQLELIGHSTDFPHPCDQELQALTPQOTLSRKKEAPFERCFSLRMSTLTSSRR 220
 QY 192 TLNLKATWVHLHCSGMRAYKRPACTSPAGSPRSPPLQCLVCEAIPLQPFHDA 249
 DB 221 TLNLKATWVHLHCSGMRAYKRPACTSPAGSPRSPPLQCLVCEALP---HFGS 274
 RESULT 9
 ABG66737
 ID ABG66737 standard; protein; 407 AA.
 AC ABG66737;
 XX 30-AUG-2002 (first entry)
 DT 30-AUG-2002 (first entry)
 XX Human novel polypeptide #72.
 DE Human; inflammatory condition; shock; sepsis; immune response; cancer;
 XX wound healing; central nervous system disease; haematopoiesis;
 XX peripheral nervous system disease; amyotrophic lateral sclerosis; tendon;
 XX myeloid cell disorder; lymphoid cell disorder; platelet disorder; bone;
 XX cartilage; ligament; nerve tissue; ulcer; osteoporosis; osteoarthritis;
 XX bone degenerative disorder; periodontal disease; reperfusion injury;
 XX lung fibrosis; liver fibrosis; autoimmune disorder; bacterial infection;
 XX allergic condition; thrombolysis; thrombosis; coagulation disorder;
 XX fungal infection.
 OS Homo sapiens.
 XX WO200244340-A2.
 PN 06-JUN-2002.
 PD 30-NOV-2001; 2001WO-US047004.
 PF 30-NOV-2001; 2000US-00728952.
 XX (HYSE-) HYSEQ INC.
 PA Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Wang J, Wang D;
 PI Yamazaki V, Ujwal ML, Drmanac RT;
 DR WPI; 2002-508509/54.
 DR N-PSDB; ABK94961.
 XX Novel nucleic acids and polypeptides for diagnosis, treatment of
 PT inflammatory, autoimmune, nervous system, myeloid or lymphoid cell
 PT disorders, cancer and promoting wound healing.
 XX Claim 10; Page 649; 672pp; English.
 CC The invention relates to human novel polynucleotides and associated
 CC polypeptides. The polynucleotides and polypeptides are useful for
 CC treating inflammatory conditions such as arthritis, nephritis, Crohn's
 CC disease, ischaemia-reperfusion injury, shock, sepsis, immune responses
 CC and cancer and for promoting wound healing. The sequences are used to
 CC induce the proliferation of neural cells and regeneration of nerve and
 CC brain tissue, and are useful for the treatment of central and peripheral
 CC nervous system diseases and neuropathies, such as Alzheimer's disease,
 CC Parkinson's disease, Huntington's disease and amyotrophic lateral
 CC sclerosis. The sequences are involved in chemotactic or chemokinetic
 CC activity, regulation of haematopoiesis, treatment of myeloid or lymphoid
 CC cell disorders and platelet disorders such as thrombocytopenia,
 CC regeneration of bone, cartilage, tendon, ligament and/or nerve tissue

CC growth, tissue repair, healing of burns, incisions, ulcers, treatment of
 CC osteoporosis, osteoarthritis, bone degenerative disorders and periodontal
 CC disease. The sequences of the invention are also useful for gut
 CC protection or regeneration and treatment of lung or liver fibrosis.
 CC reperfusion injury in various tissues, immune deficiencies and disorders
 CC including severe combined immunodeficiency (SCID), bacterial or fungal
 CC infections, autoimmune disorders e.g. multiple sclerosis and myasthenia
 CC gravis, allergic conditions such as asthma, thrombolytic or thrombotic
 CC and coagulation disorders. Sequences ABG6666-ABG66758 represent human
 CC novel polypeptides of the invention
 XX
 SQ Sequence 407 AA;

Query Match 48.7%; Score 797; DB 5; Length 407;
 Best Local Similarity 67.9%; Pred. No. 2.4e-73;
 Matches 171; Conservative 11; Mismatches 28; Indels 42; Gaps 5;

QY 10 SNTLEKREKSRDARSRSQETEVLYQLAHTLPFARGVSAHLDKASIMRLTISYLMHR 69
 DB 54 STTLEKREKSRDARSRSQETEVLYQLAHTLPFARGVSAHLDKASIMRLTISYLMHR 113
 QY 70 CAAGKRGATGRLPEGGGFRHGTTHRRGRLPYGKCQ-----QAPGQSV 118
 DB 114 CAAG-----EWNQVAGGGEPLDACYLKALGFWVWLTAEQDMAY- 152
 QY 119 LCSSSLIHNPTPGTNFS-LELIGHSTFDPIHPCDOBELQDALTPRPNLSSKKLEAPTERH 177
 DB 153 -----LSENVSKHGLSQLELIGHSTFDPIHPCDOBELQDALTPQQLTSRRKVAEPIERC 207
 QY 178 FSLRMKSTLTSGRTLNKAAATWVKVLCGSHMRAYKPPAQTSPAGSPREPPLOCLVLI 237
 DB 208 FSLRMKSTLTSGRTLNKAAATWVKVLCGSHMRAYKPPAQTSPAGSPREPPLOCLVLI 267
 QY 238 EAIPLPFTDGA 249
 DB 268 EAIPLPFTDGA 275

RESULT 10

AAB34826
 ID AAB34826 standard; protein; 705 AA.

XX AAB34826;

XX 28-MAY-2003 (first entry)

XX Human hypoxia inducible factor (HIF)-3alpha protein.

XX Human; hypoxia inducible factor; HIF-3alpha; tumour; muscular function;
 KW vascular endothelial growth factor; VEGF; angiogenesis; cardiovascular;
 KW coronary disease; cardiac disease; cytostatic; cardiac; cell survival.

XX Homo sapiens.

XX MO20294862-A2.

XX 28-NOV-2002.

XX 23-MAY-2002; 2002MO-CA000752.

XX 23-MAY-2001; 2001US-0292630P.

XX 08-FEB-2002; 2002US-0354529P.

XX (ANGI-) ANGIOGENE INC.

XX Guy L;

XX WPI; 2003-129410/12.

XX N-PSDB; AAD53131.

XX New Hypoxia Inducible Factor-3 alpha polynucleotide for inducing vascular
 PT endothelial growth factor (VEGF) expressions in a tissue or cell,
 PT particularly for eliminating tumors, or treating coronary or cardiac

PT diseases in mammals.

XX Claim 13; Col 69-72; 38pp; English.

XX The invention relates to hypoxia inducible factor (HIF)-3alpha DNA and
 CC protein sequences. The HIF-3alpha DNA and protein sequences are useful
 CC for inducing vascular endothelial growth factor (VEGF) expression or
 CC angiogenesis in a mammalian tissue or cell, for modulating tumoural cell
 CC survival, eliminating a tumoural cell, or evaluating the malignancy of a
 CC tumour in a subject. They are also useful for improving muscular
 CC function, particularly for treating coronary and cardiac diseases in
 CC mammals. The present sequence is human HIF-3alpha protein

XX Sequence 705 AA;

Query Match 48.6%; Score 795; DB 6; Length 705;
 Best Local Similarity 67.6%; Pred. No. 8.6e-73;
 Matches 171; Conservative 11; Mismatches 29; Indels 42; Gaps 5;

QY 9 RSNTLEKREKSRDARSRSQETEVLYQLAHTLPFARGVSAHLDKASIMRLTISYLMHR 68
 DB 3 STTLEKREKSRDARSRSQETEVLYQLAHTLPFARGVSAHLDKASIMRLTISYLMHR 62
 QY 69 LCAAGKRGATGRLPEGGGFRHGTTHRRGRLPYGKCQ-----QAPGQSV 117
 DB 63 LCAAG-----EWNQVAGGGEPLDACYLKALGFWVWLTAEQDMAY 102
 QY 118 LCSSSLIHNPTPGTNFS-LELIGHSTFDPIHPCDOBELQDALTPRPNLSSKKLEAPTERH 176
 DB 103 -----LSENVSKHGLSQLELIGHSTFDPIHPCDOBELQDALTPQQLTSRRKVAEPIERC 156
 QY 177 FSLRMKSTLTSGRTLNKAAATWVKVLCGSHMRAYKPPAQTSPAGSPREPPLOCLVLI 236
 DB 157 FSLRMKSTLTSGRTLNKAAATWVKVLCGSHMRAYKPPAQTSPAGSPREPPLOCLVLI 216
 QY 237 CEAIPLPFTDGA 249
 DB 217 CEAIPLPFTDGA 225

RESULT 11

AAB94934
 ID AAB94934 standard; protein; 648 AA.

XX AAB94934;

XX 26-JUN-2001 (first entry)

XX Human protein sequence SEQ ID NO:16415.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUN-2000; 2000EP-00116126.

XX 29-JUN-1999; 99JP-00248036.

XX 27-AUG-1999; 99JP-00300253.

XX 11-JAN-2000; 2000JP-00118776.

XX 02-MAY-2000; 2000JP-00183767.

XX 09-JUN-2000; 2000JP-002441899.

XX (HELI-) HELIX RES INST.

XX Oca T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI; 2001-318749/34.

PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

PS Claim 8; SEQ ID NO 16415; 2537bp + Sequence Listing; English.

XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention

CC Sequence 648 AA;

XX

SO Query Match 35.0%; Score 573; DB 4; Length 648; Best Local Similarity 61.0%; Pred. No. 7.7e-50; Matches 125; Conservative 11; Mismatches 27; Indels 42; Gaps 5;

QY 57 MRLTSLVLRMRRLCAAGKGRARGLRLLPEGPGFRGTRHGRHGVKGCQ----- 109
1 MRLTSLVLRMRRLCAAG-----EMQVAGAGBPLDACYKALEGF 40
DB

QY 110 -----QAPGPOSVDCSSSLINHPPTGTFNS-LELIGHSIFPFIRPCQBELQDALTPRPN 164
41 VMVLTLAGDMAY-----LSENVSKHGLSGLELIGHSIFPFIRPCQBELQDALTPRQT 94
DB

QY 165 LSKKKLEAPTRRHRFSLRMKSTLTISRGRITLNKAAATWKLHSCGMRAYKPPAQTSPPAGSP 224
95 LSRKRVAPTRCPSLRMKSTLTISRGRITLNKAAATWKLHSCGMRAYKPPAQTSPPAGSP 154
DB

QY 225 RSEBPLQCLVLICAIIPQLPFRHDA 249
155 DSEBPLQCLVLICAIIP---HPGS 175
DB

RESULT 12
AAB93710 ID AAB93710 standard; protein; 199 AA.
XX
AC AAB93710;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:13303.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS Homo sapiens.
XX
PN EPI074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX

PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INSTR.
XX
PI Oka T, Isegai T, Mshikawa T, Hayaishi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

PS Claim 8; SEQ ID NO 13303; 2537bp + Sequence Listing; English.

XX

XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention

CC Sequence 199 AA;

XX

SO Query Match 31.7%; Score 518.5; DB 4; Length 199; Best Local Similarity 68.0%; Pred. No. 6.2e-45; Matches 104; Conservative 9; Mismatches 13; Indels 27; Gaps 2;

QY 120 CSSSLINHPPTGTFN-----SLELIGHSIFPFIRPCDDEBQ 156
14 CTSWLTTRCPSPASAPATWRPLSCAPSATCACTASAPQLELIGHSIFPFIRPCDDEBQ 73
DB

QY 157 DALTPRPNSKKKLEAPTRRHRFSLRMKSTLTISRGRITLNKAAATWKLHSCGMRAYKPPA 216
74 DALTPRQTLSSRKRVAPTRCPSLRMKSTLTISRGRITLNKAAATWKLHSCGMRAYKPPA 133
DB

QY 217 QTSPPAGSPRSEBPLQCLVLICAIIPQLPFRHDA 249
134 QTSPPAGSPSEBPLQCLVLICAIIP---HPGS 162
DB

RESULT 13
AD116229 ID AD116229 standard; protein; 513 AA.
XX
AC AD116229;
XX
DT 22-APR-2004 (first entry)
XX
DE Human nucleic acid-associated protein (NAAP) #14.
XX
KW human; nucleic acid-associated protein; NAAP; autoimmune disorder;

